Hybrid semantics for Bio-PEPA

Hybrid semantics for Bio-PEPA

Vashti Galpin

Laboratory for Foundations of Computer Science, School of Informatics
University of Edinburgh, United Kingdom

Abstract

This paper investigates the stochastic, continuous and instantaneous (hybrid) modelling of systems defined in Bio-PEPA, a quantitative process algebra for biological modelling. This is achieved by mapping a Bio-PEPA model to a model in stochastic HYPE, a process algebra that models these three behaviour types in a compositional and structured manner. The novel mapping between process algebras provides another method of analysis for Bio-PEPA models and presents the modeller with a well-structured stochastic HYPE model which can itself be easily modified and is only a small constant larger in size than the Bio-PEPA model. The structure of the stochastic HYPE model generated has desirable properties and also gives a general framework for modelling biochemical systems where the advantages of both stochastic and deterministic simulation are required. Thresholds are introduced for each reaction, and when all values are above these thresholds, the reaction is treated deterministically. However, if a relevant value is below a threshold, the reaction is treated stochastically (as are the changes in species quantities as a result of that reaction). It is proved that in the purely deterministic case and in the purely stochastic case, the stochastic HYPE model has the same behaviour as the Bio-PEPA model when considered purely deterministically and purely stochastically, respectively. Furthermore, addition of instantaneous events in the style of Bio-PEPA with events is illustrated, and a proposal for mapping Bio-PEPA with delays (Bio-PEPAd) to stochastic HYPE is presented.

Keywords: process algebra, biological modelling, events, hybrid, stochastic, deterministic, continuous, discrete, instantaneous, simulation

1. Introduction

This paper presents a mapping from Bio-PEPA [30] to stochastic HYPE [7, 9] with the outcome that the system represented by a Bio-PEPA model can be analysed in a stochastic and hybrid (continuous and instantaneous) fashion, allowing for both continuous and stochastic modelling of reactions, switching dynamically between the different behaviours when appropriate, by the use of instantaneous transitions. This approach is novel, as the mapping provides a well-structured process algebra model that can be further modified, for example, by added additional time-based or threshold-based events in the style of Bio-PEPA with events [26]. The mapping in effect defines a general framework for modelling biochemical reaction systems in a stochastic hybrid manner using the process algebra stochastic HYPE. Additionally, this approach provides another type of analysis for Bio-PEPA models, on top of the existing approaches which include various types of simulation, model checking and continuous-time Markov chain (CTMC) analysis.

Traditionally, simulation of models in system biology has been carried out either through the numerical solution of ordinary differential equations (ODEs) that define a generalised mass action (GMA) model, or through stochastic simulation giving a realisation of a chemical master equation model (CME) using Gillespie's algorithm [42], or an exact or approximate variant thereof [43, 20, 18]. The ODE method treats quantities (either molecule counts or concentrations) as *continuous* and provides a single *deterministic* trace of the behaviour of the system. It has the benefit of being fast (although stiff systems can present problems and require the use of implicit algorithms for good performance). The *stochastic* method treats each reaction as *discrete* and each trace is a single possible realisation of

Email address: Vashti.Galpin@ed.ac.uk (Vashti Galpin)

the behaviour of the system. To obtain an idea of average behaviour, multiple traces must be obtained. This approach is closer to our understanding of biochemistry than the ODE approach but can be computationally expensive.

The low cost of ODE approaches has led to this approach being dominant in past decades. However, as computer power has increased and become more parallel, stochastic simulation is now frequently used. It is an example of an *embarrassingly parallel* problem as each simulation can be executed independently of another, hence allowing multiple simulation runs to be spread over many processors.

Hybrid approaches have also been proposed as a way to mitigate the cost of stochastic simulation where part of the simulation is run stochastically and part is run by other faster methods including ODEs. Some of these approaches are static in nature in that the identification of the part of the system that will be run stochastically is done in advance of simulation. Others are dynamic and the simulation algorithm switches between stochastic and other approaches as appropriate, defined by thresholds on the values of molecule numbers, reaction speeds or propensities. Examples of these are discussed in two recent survey papers [36, 58].

A recent development in systems biology is the use of formal languages to describe reaction systems. These languages or methods were originally developed in computer science for the specification, verification and performance modelling of human-created systems such as networks and computer systems. The languages used for biological modelling are translated into the mathematical structures of interest: ODEs for input into an appropriate solver or input for a general stochastic simulation program. One of the advantages of these languages is the ability to provide an unambiguous description which is separate from programs that perform analysis and simulation. Examples are κ-calculus [33], π-calculus [63, 61, 6], Beta-binders [60], Bio-Ambients [62], sCCP [11], continuous π-calculus [54] and LBS [59].

This paper focusses on Bio-PEPA [30] which was developed from the stochastic process algebra PEPA [48]. Bio-PEPA models can be analysed in a number of ways including stochastic simulation and deterministic simulation, but it has not been possible to analyse it using a combination of these methods. This analysis will be provided by mapping a Bio-PEPA model to a stochastic HYPE model. Stochastic HYPE is a process algebra encompassing instantaneous, stochastic and continuous behaviour[1], and whose semantics are defined by transition-driven stochastic hybrid automata (TDSHA) [13]. The decision to map to stochastic HYPE rather than its semantics is due to the fact that the mapping provides a well-structured language-based model which is easy to modify and also provides a general framework for describing reaction systems as stochastic hybrid models. When a Bio-PEPA model is expressed in stochastic HYPE, it is possible to treat each reaction (and changes to quantities of species involved in the reaction) stochastically or deterministically. The treatment of a reaction is determined by thresholds on reactant species and the reaction rate. If any value is below its threshold, the reaction will be treated stochastically and if all values are above their thresholds, the reaction and the species will be treated deterministically. Note that a species can be treated stochastically in one reaction and deterministically in another, thus allowing for stochastic treatment only when necessary.

The structure of this paper is as follows. In the next section, the advantages and the disadvantages of the continuous deterministic and the stochastic discrete approaches are considered, motivating the combination of the two approaches. After that Bio-PEPA is introduced with an example which will be used through this paper to illustrate various notions. Stochastic HYPE and the mapping of Bio-PEPA to stochastic HYPE is given followed by the properties of the HYPE model. The next section is a case study after which a discussion of the general approach appears. Next, the addition of events in the style of Bio-PEPA with events and a proposal to extend HYPE to allow the mapping of Bio-PEPA with delays are described and finally related work, further work and conclusions are presented.

2. Deterministic, stochastic and inbetween

Recent articles have surveyed the different approaches that are available for deterministic, stochastic and hybrid modelling [58, 36] and give guidelines for choosing between the two main approaches of deterministic and stochastic [64, 70]. Criteria include the specific objective of the model, limitations in terms of computational power and time, availability of experimental data, and the level of detail or accuracy required. As usual, Box's statement applies, "all models are wrong but some are useful" [15].

[1] The process algebra HYPE models deterministic and instantaneous behaviour [40]. Stochastic HYPE is an extension of HYPE that additionally models stochastic behaviour.

Assuming a spatially uniform mixture of species, at the thermodynamic limit (where the size of the volume and number of molecules goes to infinity and the density/concentration remains constant), the stochastic model and deterministic model have the same behaviour and the expectation of the stochastic model is the solution of the deterministic model as shown by Kurtz [52]. However, since finite volumes are usual in this type of modelling, discrepancies are often found between the stochastic model and the deterministic model.

An example of this is Keizer's Paradox [50] as analysed in [69]. In the autocatalytic model studied, it is shown that as time increases, the results of the two approaches differ. The deterministic model has a non-zero steady state (it has two steady states: one zero and unstable, and one non-zero and stable) and in the stochastic model, the species of interest is eventually depleted (its only steady state is the zero state representing a fixed distribution). However, it is not possible to exchange increasing time (the steady states occur as time goes to infinity) for increasing volume in the results of Kurtz [52] mentioned above and hence this is not a violation of these results. This example does demonstrate that for models with limited volumes the two different modelling techniques can give different outcomes as time tends to infinity.

Stochastic traces can also exhibit quasicycles [57] which show as oscillations in the output although the deterministic instantiation of the model has no limit cycle and may show at most damped oscillations before a steady state. Sometimes in the case of models where there are quasicycles, there are limit cycles for some combination of parameters, suggesting that the model does somehow contain oscillatory behaviour. However, it has been shown that there are models without limit cycles over the whole parameter space that can generate quasicycles [55]. Exclusion of these models on the grounds that their deterministic instantiation shows no oscillatory behaviour may incorrectly exclude mechanisms that do work in a stochastic reality.

An example of the presence of quasicycles is in the stochastic modelling of the *Ostreococcus tauri* circadian clock [2, 67]. In a particular lighting condition, that of constant light, the ODE simulation displayed damping oscillations leading to a steady state, whereas the stochastic simulation of a single cell showed continued oscillations. The ODE trace matched that of the experimental data. However, the data was from a population of cells, hence providing an average behaviour and a reasonable hypothesis is that due to the lack of day-night cycle, the cells were demonstrating individual oscillating behaviour but were becoming unsynchronised in their behaviour, hence leading to a steady state average behaviour. This indicates that the stochastic simulation could be an accurate model of a single cell, but as yet there is no current single-cell data to support this. This is an example of why it is necessary to understand the potential differences between deterministic and stochastic modelling output.

It has been argued that in terms of accuracy, for low molecule counts, stochastic simulation should be used. A counter to this argument is that generally, multiple stochastic simulation traces are required to obtain an idea of average behaviour and that this resembles the output of the deterministic trace [70]. Additionally, with single-cell experiments, multiple replicates are necessary to reduce random variation and hence multiple stochastic simulations are also required. However, in the case of circadian clocks as described above, if single cell data were available, it would appear that a simple averaging approach would not be useful, showing the importance of understanding the differences between the two modelling approaches.

It is clear from the arguments above that a deterministic approach can differ from a stochastic approach, for example by hiding details of fluctuations that a stochastic approach might reveal, and hence there is a strong case for stochastic modelling based on accuracy. On the other hand, stochastic modelling can be expensive in terms of computing power and time, hence there is a practical reason for using deterministic modelling when appropriate.

Considering the practical issues, a further reason for using deterministic models is to reduce the cost of modelling when there are reactions on different time scales. ODE systems are called stiff when the rates of reactions vary by orders of magnitude [32, 17]. This definition is also used for stiff systems in a stochastic context [18]. When these systems are simulated using an exact method, the fast reactions are frequent and result in only small advances in the time of the simulation, leading to slow performance. Variants of the Gillespie algorithm have been developed to mitigate against the effect of frequent reactions. Two basic approaches are taken: in the first, called tau-leaping, under an assumption of unchanging propensities for a specific time period, multiple reactions are simulated in that time period rather than a single time advance to the next reaction [43]. In the second approach, a form of quasi-steady state assumption or quasi-equilibrium approximation is used (similar to that applied in the derivation of the Michaelis-Menten reduction [56]) where the fast reactions are assumed to lead quickly to an equilibrium amongst some species in the system, and this equilibrium is used to avoid simulation of the fast reactions [20, 18]. Hybrid approaches can also be used to simulate fast reactions with faster approaches such as deterministic approximation of

ODEs [45, 49, 68].

Considering all of these issues, it appears that a mixture of stochastic and deterministic simulation solves some of the problems, leading to relatively fast simulation while still using stochastic methods when necessary to identify or reveal fluctuations, quasicycles or low copy count activity. This, together with the goal of providing a process algebra hybrid model, is the motivation for the research in the following sections.

3. Bio-PEPA

This section presents Bio-PEPA [30] in a moderately informal manner[2], providing sufficient detail to support the definition of the translation to stochastic HYPE and is illustrated with a small example. The main components of a Bio-PEPA system are the sequential or species components describing the behaviour of each of the biochemical species, and the model component which combines the species components and hence models the interactions between the species. Additionally, a context is defined to store information such as functional rates, compartments and parameters.

The syntax of the sequential/species components is defined by the grammar $S ::= (\alpha, \kappa)\ \mathtt{op}\ S \mid S + S \mid C$ with $\mathtt{op} ::= \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$. Constants, which are species names, are defined to be a specific sequential component using the notation $C \stackrel{def}{=} S$. A **well-defined Bio-PEPA species** C has the form

$$C \stackrel{def}{=} (\alpha_1, \kappa_1)\ \mathtt{op}_1\ C + \ldots + (\alpha_n, \kappa_n)\ \mathtt{op}_n\ C \quad \text{(written as} \quad C \stackrel{def}{=} \sum_{i=1}^{n} (\alpha_i, \kappa_i)\ \mathtt{op}_i\ C)$$

where $\alpha_i \neq \alpha_j$ for $i \neq j$. In each prefix term $(\alpha_i, \kappa_i)\ \mathtt{op}_i\ C$, α_i is an action name from a set $\mathcal{A}$ and gives the name or label of a reaction, κ_i is the stoichiometric coefficient[3] of the species and the prefix combinator $\mathtt{op}_i$ represents the role of the element in the reaction. If a species is a reactant in the reaction then $\downarrow$ is used, if a product then $\uparrow$, if an activator then $\oplus$, if an inhibitor then $\ominus$, and $\odot$ is used for a generic modifier[4]. The operator + expresses the choice between two sequential components.

The grammar for model components is $Q ::= Q \underset{L}{\bowtie} Q \mid S(x)$. A **well-defined Bio-PEPA model** P has the form

$$P \stackrel{def}{=} C_1(x_1) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} C_p(x_p), \quad \text{(written as } P \stackrel{def}{=} \underset{*}{\bowtie}_{i=1}^{p} C_i(x_i))$$

where each C_i is a well-defined species and if $i \neq j$ then $C_i \neq C_j$. The use of $\underset{*}{\bowtie}$ indicates all shared reactions are synchronised on. A well-defined model is the synchronisation of well-defined species on shared reactions. In the model component $C(x)$, the parameter $x \in \mathbb{R}$ represents the molecular count for that species[5].

Additional information is required, including reaction rate equations, and this gives rise to a **Bio-PEPA system** which is a 6-tuple $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P \rangle$, where $\mathcal{V}$ is the set of compartments, $\mathcal{N}$ is the set of quantities describing each species, $\mathcal{K}$ is the set of parameters, $\mathcal{F}$ is the set of functional rates, $Comp$ is the set of well-defined sequential components and P is a well-defined model component.

3.1. Example: enzyme kinetics

A small running example focussing on enzyme kinetics is now introduced. Enzymes are proteins that catalyse reactions. Typically, a substrate (S) binds to an enzyme (E) to form a complex (C). The complex can dissociate into the substrate and the enzyme. Alternatively, the catalytic reaction can happen and the complex dissociates into a product (P) and the enzyme. These three reactions are described in Figure 1 together with the Bio-PEPA species

[2]The reader is referred to the journal paper [30] and the associated technical report [29] for full details.

[3]The stoichiometry/stoichiometric coefficient of a species with respect to a specific reaction is the relative quantity of that species involved in the reaction compared to other species in the reaction. At the molecule level, it describes the exact number of molecules. In the reaction $A + 3B \rightarrow 2C + D$, three times as much B as A is consumed to produce as much D as A and twice as much C as A.

[4]The generic modifier is used when a species plays a role in a reaction that is different to any of the other roles. This has been useful for constructing Bio-PEPA models that abstract from a reaction-based view and also applying Bio-PEPA in non-biological contexts.

[5]Quantities other than molecule counts can be used, for example, molar concentrations. Converting between concentrations and counts involves the Avogadro constant and the volume. Reaction rates may need to be scaled. For a more detailed explanation, see [30, 29]. In the current paper, only molecule counts are used.

$$\begin{array}{rcll} S + E & \rightarrow & C & f_1 = k_1 S E \\ C & \rightarrow & S + E & f_2 = k_2 C \\ C & \rightarrow & P + E & f_3 = k_3 C \end{array}$$

$$\begin{array}{rcl} S & \stackrel{def}{=} & (\alpha_1, 1) \downarrow S + (\alpha_2, 1) \uparrow S \\ E & \stackrel{def}{=} & (\alpha_1, 1) \downarrow E + (\alpha_2, 1) \uparrow E + (\alpha_3, 1) \uparrow E \\ C & \stackrel{def}{=} & (\alpha_1, 1) \uparrow E + (\alpha_2, 1) \downarrow E + (\alpha_3, 1) \downarrow E \\ P & \stackrel{def}{=} & (\alpha_3, 1) \uparrow P \end{array}$$

Figure 1: Reactions and rates for the substrate-enzyme-product example (left) and Bio-PEPA species for the substrate-enzyme-product (right)

$$\texttt{prefixReac}\ \frac{}{(\alpha,\kappa)\downarrow S(l) \xrightarrow{(\alpha,[S:\downarrow(l,\kappa)])}_c S(l-\kappa)}\ \kappa \leq l \leq N_S \qquad \texttt{prefixProd}\ \frac{}{(\alpha,\kappa)\uparrow S(l) \xrightarrow{(\alpha,[S:\uparrow(l,\kappa)])}_c S(l+\kappa)}\ 0 \leq l \leq N_S - \kappa$$

$$\texttt{prefixMod}\ \frac{}{(\alpha,\kappa)\ \texttt{op}\ S(l) \xrightarrow{(\alpha,[S:\texttt{op}(l,\kappa)])}_c S(l)}\ \begin{array}{l}\kappa \leq l \leq N_S \text{ if } \texttt{op} = \oplus \\ 0 \leq l \leq N_S \text{ if } \texttt{op} \in \{\ominus, \odot\}\end{array}$$

$$\texttt{choice1}\ \frac{S_1(l) \xrightarrow{(\alpha,w)}_c S_1'(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,w)}_c S_1'(l')} \qquad \texttt{choice2}\ \frac{S_2(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha,w)}_c S_2'(l')}$$

$$\texttt{coop1}\ \frac{P_1 \xrightarrow{(\alpha,w)}_c P_1'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1' \underset{L}{\bowtie} P_2}\ \alpha \notin L \qquad \texttt{coop2}\ \frac{P_2 \xrightarrow{(\alpha,w)}_c P_2'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w)}_c P_1 \underset{L}{\bowtie} P_2'}\ \alpha \notin L$$

$$\texttt{coop3}\ \frac{P_1 \xrightarrow{(\alpha,w_1)}_c P_1' \quad P_2 \xrightarrow{(\alpha,w_2)}_c P_2'}{P_1 \underset{L}{\bowtie} P_2 \xrightarrow{(\alpha,w_1::w_2)}_c P_1' \underset{L}{\bowtie} P_2'}\ \alpha \in L \qquad \texttt{constant}\ \frac{S(l) \xrightarrow{(\alpha,[S:\ \texttt{op}\ (l,\kappa)])}_c S'(l')}{C(l) \xrightarrow{(\alpha,[C:\ \texttt{op}\ (l,\kappa)])}_c S'(l')}\ C \stackrel{def}{=} S$$

$$\texttt{Final}\ \frac{P \xrightarrow{(\alpha,w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P\rangle \xrightarrow{(\alpha, r_\alpha[w,\mathcal{N},\mathcal{K}])}_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P'\rangle}$$

Figure 2: Operational semantics of Bio-PEPA

definitions. This model uses mass action for all reactions, as given after each reaction on the left in Figure 1. Bio-PEPA allows mass action for reaction, as well as a wide range of other rate functions. The Bio-PEPA model is then $EK \stackrel{def}{=} S(S_0) \underset{*}{\bowtie} E(E_0) \underset{*}{\bowtie} C(C_0) \underset{*}{\bowtie} P(P_0)$ where S_0, E_0, C_0 and P_0 are the initial quantities of each species.

Note that this is a collection of bimolecular reactions (in the sense that there at most two molecules as reactants) and is a direct representation of the reality of reactions. It is also possible to consider this reaction more abstractly and view as a single reaction that transforms the substrate to the product through the involvement of the enzyme. This can be represented in Bio-PEPA as a single reaction as follows.

$$S' \stackrel{def}{=} (\beta, 1)\downarrow S' \quad E' \stackrel{def}{=} (\beta, 1) \oplus E' \quad P' \stackrel{def}{=} (\beta, 1)\uparrow P' \quad EK' \stackrel{def}{=} S'(S_0') \underset{*}{\bowtie} E'(E_0') \underset{*}{\bowtie} P'(P_0')$$

The rate for the reaction, f_β will then be defined according to Michaelis-Menten kinetics [56].

3.2. Semantics

To describe the behaviour of a Bio-PEPA system, semantics must be defined in terms of the operators of the algebra. The operational semantics for Bio-PEPA systems are given in Figure 2 where N_S is the maximum number of molecules for the species S. These operational semantics define two distinct labelled transition systems[6].

The first, the capability relation ($\rightarrow_c$), is defined as over Bio-PEPA model components and has labels of the form (a, w) with $w ::= [\ S : \texttt{op}\ (l, \kappa)\ ] \mid w :: w$ where S is the species name, l is the quantity of S and κ the stoichiometric

[6]Note that these labelled transition systems are sets, unlike PEPA where the transition system defined is a multiset.

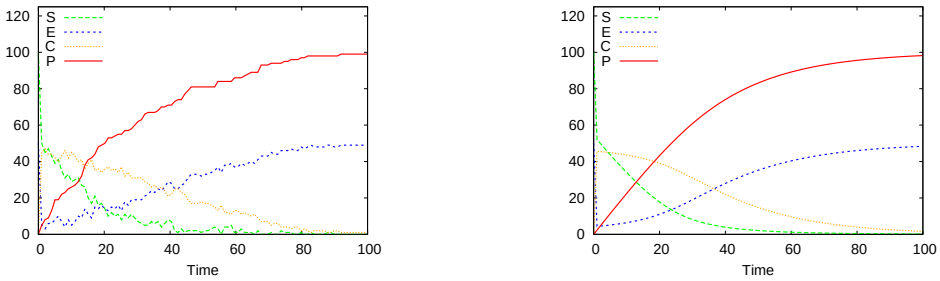

Figure 3: Stochastic simulation (left) and deterministic simulation (right) of the Bio-PEPA *EK* model with $k_1 = 50$, $k_2 = 100$, $k_3 = 0.05$, $S_0 = 100$, $E_0 = 50$ and $C_0 = P_0 = 0$

coefficient. This relation has labels that capture the information about the species that take part in a reaction. For well-defined models, there are no repeated items in the list w (which can be also be viewed as a multiset since ordering of items is irrelevant) and hence it is a set [37].

The second transition system, the stochastic relation ($\rightarrow_s$), is defined over Bio-PEPA systems and is inferred from capability relation transitions using the rule `Final`. Its transitions are labelled with the reaction name and a value which describes the exponential distribution from which the rate of reaction is drawn. This value is calculated by applying the function defining the reaction rate to current counts of the relevant species (taken from the string w) and the appropriate parameters from $\mathcal{K}$. It has been shown that for well-defined models and systems that for each reaction, at most one transition is possible [37].

An important feature of well-defined Bio-PEPA models and systems is that it is possible to use a vector-notation representation for the states of a transition system that capture the quantity of each species. A model of the form $C_1(x_1) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} C_p(x_p)$ has states of the form $(x_1, \ldots, x_p)$.

Bio-PEPA semantics can be interpreted in a number of ways, and two approaches most relevant to this paper are now described briefly. Note that these two approaches are distinct and provide two separate forms of analysis – a Bio-PEPA model can be viewed either stochastically or deterministically but not both.

The first approach is stochastic simulation using Gillespie's algorithm [42] and variants thereof [43, 20, 18]. This method is based on the CTMC obtained from the Bio-PEPA model (namely the stochastic relation) where states are representing in vector notation as described above, and a transition between a source state and a target state is a reaction which can occur when there are sufficient reactants in the source state and each value in the target state is equal to the value in the source state modified by the stoichiometric coefficient of the reaction for that species. The rate of the transition is obtained from the rate function using the relevant species quantities. Molecule counts remain integral in this approach[7].

The second approach is a deterministic approach using the ODEs obtained from the Bio-PEPA model. Here the molecule counts become continuous, as this approach is an approximation to the stochastic case. The ODE system has the form $\frac{d\mathbf{X}}{dt} = \mathbf{D} \times \mathbf{v}$ where $\mathbf{X}$ is the vector such that $\mathbf{X}^T = (X_1, \ldots, X_N)$, where each X_i is a variable for the quantity of species S_i (the same as the vector representation mentioned above), and $\mathbf{v}$ is a vector containing the kinetic laws for each reaction, so each element has the form $f_{\alpha_k}(X_1, \ldots, X_N)$. $\mathbf{D} = \{d_{i,k}\}$ is the $N \times K$ stoichiometry matrix which can be obtained from a well-defined Bio-PEPA model by considering the prefixes of the form $(\alpha_{i,j}, \kappa_{i,j})\ \mathsf{op}_{i,j}\ S_i$ for species S_i where $\alpha_{i,j}$ is the kth reaction, as follows: if $\mathsf{op}_{i,j} = \uparrow$ then $d_{i,k} = \kappa_{i,j}$ else if $\mathsf{op}_{i,j} = \downarrow$ then $d_{i,k} = -\kappa_{i,j}$ else $d_{i,k} = 0$.

Applying this to the enzyme example, the following ODEs are obtained for the four species.

$$\begin{array}{lcllcl} dS/dt & = & -k_1 S.E + k_2 C & dE/dt & = & -k_1 S.E + k_2 C + k_3 C \\ dP/dt & = & +k_3 C & dC/dt & = & +k_1 S.E - k_2 C - k_3 C \end{array}$$

[7] The prefix rules for the capability relation given in Figure 2 impose bounds, both above and below that limit the quantity of each species; thus ensuring a finite transition system. Typically, in implementations of simulation, the only bounds enforced are those imposed by stoichiometry, and hence these systems can be unbounded in terms of species quantities.

Figure 3 gives stochastic simulation (single run) and deterministic simulation output of the Bio-PEPA model $EK \stackrel{def}{=} S(100) \underset{*}{\bowtie} E(50) \underset{*}{\bowtie} C(0) \underset{*}{\bowtie} P(0)$ using the Bio-PEPA Eclipse Plugin software [66].

Bio-PEPA has also been applied to the modelling of many different biological systems including Goldbeter's model of cyclin oscillation [29, 25], the Repressilator [28], the NF-κB signalling pathway [27], the MAPK model [35], circadian clocks [1, 2], the gp130/JAK/STAT pathway [44] and trafficking of Src in the mammalian cell [38].

4. Stochastic HYPE

Stochastic HYPE provides three distinct types of behaviour: instantaneous events that happen instantly when their activation conditions become true, stochastic events that occur after a random delay, in this case specified by an exponential distribution, and continuous behaviour where the values of the variables of the system change smoothly as defined by a system of ODEs. A simulation trace of a stochastic HYPE model consists of smoothly varying curves for each variable interspersed with discontinuities or jumps when events (stochastic or instantaneous) change the ODE system in operation or modify values of the variables of the system via resets.

The behaviour of a stochastic HYPE model is determined by mapping the model to a mathematical formalism, namely a transition-driven stochastic hybrid automaton (TDSHA) [13, 14]. In choosing a mapping for Bio-PEPA to a stochastic hybrid formalism, the choice is between a language-based formalism such as stochastic HYPE and a mathematical formalism such as TDSHA. The disadvantage of using a language-based formalism such as stochastic HYPE is that it adds an additional mapping in reaching the behaviour of the model; however in this case, this disadvantage is outweighed by obtaining a structured stochastic HYPE model that is readable due to separation of concerns, easy to modify, and elucidates how one can construct a stochastic hybrid model of a reaction system directly in stochastic HYPE. Stochastic HYPE is a powerful formalism and it is not immediately clear how to express the enzyme example of the previous section in stochastic HYPE. The mapping developed in the next section will elucidate this, hence illustration of stochastic HYPE through an example will be delayed until then.

As with the Bio-PEPA presentation, the aim of the following stochastic HYPE presentation is to give sufficient details for the mapping. More details about HYPE in both its stochastic variant and non-stochastic variant can be found in [39, 40]. Note that in the rest of this paper, HYPE will be used to mean stochastic HYPE, unless otherwise specified.

Well-defined HYPE subcomponents represent the uncontrolled capabilities of the system and have the form

$$S(\mathcal{W}) \stackrel{def}{=} \sum_{j=1}^{n} \mathrm{a}_j{:}(\iota, r_j, I_j(\mathcal{W})).S(\mathcal{W}) + \underline{\mathrm{init}}{:}(\iota, r, I(\mathcal{W})).S(\mathcal{W})$$

where $\mathcal{W} \subseteq \mathcal{V} = \{V_1, \ldots, V_n\}$ a set of real variables, $\mathrm{a}_j \neq \mathrm{a}_k$ for $j \neq k$ and $\mathrm{a}_j \neq \underline{\mathrm{init}}$ for all j. Each a_j is an **event**, either **instantaneous** ($\underline{\mathrm{a}}_j \in \mathcal{E}_d$) or **stochastic** ($\overline{\mathrm{a}}_j \in \mathcal{E}_s$), and has an associated **event condition** from a set EC. Event conditions have the form $ec(\mathrm{a}) = (act(\mathrm{a}), reset(\mathrm{a}))$ where $reset(\mathrm{a})$ is a **reset** with the form $\bigwedge_{V \in \mathcal{V}} V' = \rho(V_1, \ldots, V_n)$ where V' refers to the updated variable, and $\rho : \mathbb{R}^{|\mathcal{V}|} \to \mathbb{R}$. An **activation condition** for an instantaneous event is a boolean formula with free variables in $\mathcal{V}$, while for a stochastic event, it is a function $f : \mathbb{R}^{|\mathcal{V}|} \to \mathbb{R}^+$ that describes an exponential distribution. Subcomponents also contain **influences** of the form $(\iota, r, I(\mathcal{W}))$ (from the set $\mathcal{A}$) where $\iota \in IN$ is an **influence name** and each well-defined subcomponent has exactly one influence name in all of its influences which is unique to that subcomponent. The value r describes the strength of the influence and $I(\mathcal{W}) \in IT$ describes the variables involved in the influence which allows nonlinear ODEs to be obtained. Each influence name is mapped to a variable in $\mathcal{V}$ using the function iv, and the set ID consists of a real-valued function for each influence type name $[\![I(\mathcal{W})]\!] = f(\mathcal{W})$. A subcomponent is ready to react whenever any of its events' activation condition become true or complete, after which the influence associated with that event comes into force, replacing any previous influence. By considering all the influences mapped to a particular variable for a particular configuration of the system, an ODE can be constructed using the definitions in ID to describe the evolution of that variable whenever the system is in that configuration.

A **well-defined uncontrolled system** has the form $\Sigma = S_1(\mathcal{W}_1) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} S_s(\mathcal{W}_s)$ where each subcomponent appears at most once. Furthermore subcomponents must synchronise on all shared events, hence the use of $\underset{*}{\bowtie}$. A **HYPE controller** is defined by the two-level grammar $M ::= \mathrm{a}.M \mid 0 \mid M + M$ and $Con ::= M \mid Con \underset{L}{\bowtie} Con$ with $\mathrm{a} \in \mathcal{E} = \mathcal{E}_d \cup \mathcal{E}_s$ and with $L \subseteq \mathcal{E}$. A controller does not contain influences and its role is to control and sequence event occurrences. The controller and uncontrolled system are put in cooperation to obtain the **well-defined controlled**

Prefix with influence $\dfrac{}{\langle \mathrm{a} : (\iota, r, I).E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E, \sigma[\iota \mapsto (r, I)]\rangle}\ (\mathrm{a} \in \mathcal{E})$ **Prefix without influence** $\dfrac{}{\langle \mathrm{a}.E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E, \sigma\rangle}\ (\mathrm{a} \in \mathcal{E})$

Choice $\dfrac{\langle E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \sigma'\rangle}{\langle E + F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \sigma'\rangle}\quad \dfrac{\langle F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle F', \sigma'\rangle}{\langle E + F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle F', \sigma'\rangle}$ **Constant** $\dfrac{\langle E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \sigma'\rangle}{\langle A, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \sigma'\rangle}(A \stackrel{def}{=} E)$

Cooperation without synchronisation $\dfrac{\langle E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \sigma'\rangle}{\langle E \underset{M}{\bowtie} F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E' \underset{M}{\bowtie} F, \sigma'\rangle}(\mathrm{a} \notin M)\quad \dfrac{\langle F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle F', \sigma'\rangle}{\langle E \underset{M}{\bowtie} F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E \underset{M}{\bowtie} F', \sigma'\rangle}(\mathrm{a} \notin M)$

Cooperation with synchronisation $\dfrac{\langle E, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E', \tau\rangle \quad \langle F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle F', \tau'\rangle}{\langle E \underset{M}{\bowtie} F, \sigma\rangle \xrightarrow{\mathrm{a}} \langle E' \underset{M}{\bowtie} F', \Gamma(\sigma, \tau, \tau')\rangle}(\mathrm{a} \in M, \Gamma \text{ defined})$

Figure 4: Operational semantics for HYPE

system $\Sigma \underset{*}{\bowtie} \underline{\text{init}}.Con$. This ensures the first event to occur is $\underline{\text{init}}$ which must have *true* as its activation condition, and the initial values as its resets. Additionally all events that appear in the controller must appear in the uncontrolled system. A **well-defined HYPE model** is a tuple $(ConSys, \mathcal{V}, IN, IT, \mathcal{E}, \mathcal{A}, ec, iv, EC, ID)$ where $ConSys$ is a well-defined controlled system and the other components are as defined above. In the rest of the paper, $\mathcal{V}$ is a set or tuple of variables with $\mathcal{W} \subseteq \mathcal{V}$ denoting an arbitrary subset of $\mathcal{V}$.

4.1. Semantics

An operational semantics is defined which specifies qualitatively the behaviour of a controlled system through a labelled multitransition system, similar to [48]. The labelled multitransition system is then mapped to a TDSHA to describe the quantitative behaviour of the model. In the semantics, (operational) states keep track of the current strength and influence type of each influence. This information is then used to describe the continuous behaviour of the model, while the structure of the multitransition system describes the discrete behaviour, both instantaneous and stochastic. For reasons of space, details of TDSHA are omitted from this article and the reader is referred to [13, 14].

An **operational state** of the system is a function $\sigma : IN \to (\mathbb{R} \times IT)$. The set of all operational states is $\mathcal{S}$. A **configuration** consists of a controlled system together with an operational state $\langle ConSys, \sigma\rangle$ and the set of configurations is $\mathcal{F}$. The operational semantics give a labelled multitransition system over configurations $(\mathcal{F}, \mathcal{E}, \to)$ with $\to \subseteq \mathcal{F} \times \mathcal{E} \times \mathcal{F}$ and are given in Figure 4. The only rules which modify the state are Prefix with influence and Cooperation with synchronisation. The **updating function** $\sigma[\iota \mapsto (r, I)]$ is defined as in the standard manner as $\sigma[\iota \mapsto (r, I)](x) = (r, I)$ if $x = \iota$, otherwise $\sigma(x)$. For Cooperation with synchronisation, consistency must be ensured in the way influences are updated by the cooperating components. The **change-detecting function** Γ does this by comparing the previous operational state with the new operational states. The function returns the state which differs from the previous state. However, if both new states differ from the previous state the function is not defined. $(\Gamma(\sigma, \tau, \tau'))(\iota) = \tau(\iota)$ if $\sigma(\iota) = \tau'(\iota)$ and $(\Gamma(\sigma, \tau, \tau'))(\iota) = \tau'(\iota)$ if $\sigma(\iota) = \tau(\iota)$, otherwise it is undefined. For well-defined HYPE models, the cooperation rule always succeeds [40].

An informal presentation is now given of the behaviour of a HYPE model, once mapped to a TDSHA.

Deterministic continuous behaviour Each $\langle P, \sigma\rangle$ in the labelled multitransition system becomes a mode in the TDSHA. The continuous behaviour in that mode is specified by the following set of ODEs.

$$P_\sigma = \left\{ \frac{dV}{dt} = \sum \{ r[\![I(\mathcal{W})]\!] \;\middle|\; iv(\iota) = V \text{ and } \sigma(\iota) = (r, I(\mathcal{W})) \} \;\middle|\; V \in \mathcal{V} \right\}$$

The ODE for a variable V represents all influences being applied additively, and nonlinear ODEs are possible through the use of the influence type; for example, the model can have the definition $[\![I(V_1, V_2)]\!] = V_1 V_2$ and then the rate of change of a variable V will be, in part, determined by the product of these two variables.

Stochastic discrete behaviour A transition labelled with a stochastic event $\langle P, \sigma \rangle \xrightarrow{\overline{\mathrm{a}}} \langle P', \sigma' \rangle$ is mapped to a stochastic transition (with a *true* guard) between two modes in the stochastic hybrid automata where the rate of transition is determined by the function $act(\overline{\mathrm{a}})$ and the resets by $reset(\overline{\mathrm{a}})$. The actual rate of transition between the two modes is determined by the sum of the rates of all transitions between these two modes labelled with $\overline{\mathrm{a}}$, and the probability of the transition is determined by the transition rate divided by the summed rate.

Instantaneous discrete behaviour An instantaneous transition $\langle P, \sigma \rangle \xrightarrow{\underline{\mathrm{a}}} \langle P', \sigma' \rangle$ is mapped to an instantaneous transition between two modes in the stochastic hybrid automata where the guard of the transition is determined by the boolean formula $act(\underline{\mathrm{a}})$ and the resets by $reset(\underline{\mathrm{a}})$.

This behaviour can be summarised as follows. Determined by the current mode, each variable changes continuously according to the ODE associated with that mode. Transitions to a different mode occurs either when a guard of an instantaneous transition becomes true, or when a stochastic duration has been completed. In the case of both an instantaneous transition and a stochastic transition, variable values may be modified by the resets associated with the transition. The variables then begin varying in accordance with the ODEs of the new mode.

Additionally, the focus here is on models that can never execute an infinite number of simultaneous instantaneous events called instantaneous Zeno behaviour. This ensures that models can be simulated, and that HYPE models can be interpreted as piecewise deterministic Markov processes (PDMPs) [34] for which instantaneous Zeno behaviour is not permitted. The focus has been on instantaneous Zeno behaviour because it can be determined from the syntax of the model. Other undesirable behaviour such as Zeno behaviour is more difficult to identify through static analysis.

Definition 1. A HYPE model P is well-behaved if it has a finite number of finite sequences of simultaneous instantaneous events and these sequences are independent of the initial state of the system.

To ensure well-behavedness, it is sufficient to show that the instantaneous activation graph, or I-graph, of a HYPE model is acyclic. This graph is constructed by considering the instantaneous transitions of the labelled transition systems obtained from the controller of the HYPE model. The reader is referred to [40, 7] for further information about well-behavedness and results about controllers that are well-behaved. The following two results (which have appeared previously) straightforward to prove.

Proposition 1 ([40, 9]). *Let P be a HYPE model with $Con \stackrel{\text{def}}{=} \mathrm{a}_1 \ldots \mathrm{a}_n.Con$. P is well-behaved if one of the following holds.*

1. *At least one of the a_i is stochastic.*
2. *All a_i are instantaneous, the resets for all a_i are the identity and there is no overlap in set of the values that make the guards true for each a_i.*
3. *All a_i are instantaneous and for at least one of a_i, its reset values are disjoint from the values that make the guard true for a_{i+1} (where addition is modulo n).*

Proposition 2 ([40]). *Let Con and Con' be two controllers such that for all $\underline{\mathrm{a}} \in \mathsf{ev}(Con) \setminus \mathsf{ev}(Con')$ and for $\underline{\mathrm{a}}' \in \mathsf{ev}(Con') \setminus \mathsf{ev}(Con)$, no $\underline{\mathrm{a}}$ activates an $\underline{\mathrm{a}}'$ and no $\underline{\mathrm{a}}'$ activates an $\underline{\mathrm{a}}$. If Con and Con' are well-behaved then $Con \underset{*}{\bowtie} Con'$ is well-behaved.*

5. Switching between deterministic and stochastic behaviour

Now that both formalisms of relevance to this research have been introduced it is possible to proceed with the mapping from Bio-PEPA to HYPE that will allow switching between deterministic and stochastic behaviour. This will be achieved by introducing instantaneous events that are triggered by molecule counts for a species in a reaction or rates of the reaction and which will then switch to a mode of a system where the reaction is treated differently. As noted before, the choice has been made to map to HYPE rather than its mathematical semantic objects, so that a well-structured and modifiable model is obtained.

The mapping from a Bio-PEPA model to a HYPE model will now be defined. It is useful to fix notation for the remainder of this paper. Assume a well-defined Bio-PEPA system with K reactions $\{\beta_1, \ldots, \beta_K\}$ and N species of the form

$$S_i \stackrel{def}{=} \sum_{j=1}^{n_i} (\alpha_{i,j}, \kappa_{i,j}) \ \mathsf{op}_{i,j} \ S_i$$

with each $\alpha_{i,j} \in \{\beta_1, \ldots, \beta_K\}$. To refer to the stoichiometric coefficient and operator for a specific species in a specific reaction, let $\lambda_{i,k} = \kappa_{i,j}$ when $\alpha_{i,j} = \beta_k$. This defines $\lambda_{i,k}$ to be the stoichiometric coefficient for species i in reaction k. Furthermore, f_k will be used to indicate the rate function for reaction k. The Bio-PEPA model is $S_1(x_{1,0}) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} S_N(x_{N,0})$ where each $x_{i,0}$ defines the initial quantity of the species S_i. This section will present the formal mapping and illustrate it with the actual mapping of the species E and the first reaction from the enzyme example given earlier. The complete formal mapping is given in Appendix A and the complete mapping of the enzyme example is given in Figure 6.

For the dynamic switching between stochastic and dynamic treatment of reactions to be part of the mapping, threshold values are needed. For generality, these are defined for both molecule quantities and reaction rates. In the sequel, it is assumed that each species has a threshold of $h_{i,k}$ which determines the count at which switching will occur for reaction k in which it is a reactant; at this level or below, the system will switch to treating the species (and all other species that take part in that reaction) stochastically with respect to that reaction, which means stopping continuous flows for that reaction. It is also assumed that for each reaction β_k, there is a rate r_k at which switching will occur; at this rate or below, the reaction will be considered as slow and will be treated stochastically,. For the former, a simple comparison of S_i against $h_{i,k}$ will be performed. The latter requires a rate calculation of the form $f_k(S_1, \ldots, S_N)$ to be compared with r_k. It is assumed that there are variables $X_1, \ldots, X_N$ with each X_i representing the quantity of species S_i and $\mathbf{X} = (X_1, \ldots, X_N)$.

First the species are defined as HYPE subcomponents to describe their ability to undergo continuous change under deterministic treatment of each reaction. This is similar to constructing the ODEs for a Bio-PEPA model. For each prefix $(\beta_k, \lambda_{i,k})$ in S_i which represents species S_i taking a role in reaction k, a subcomponent of the following form is created whenever S_i is a product or reactant in that reaction.

$$\begin{aligned} SC_{i,k}(\mathbf{X}) &= \underline{\text{init}} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \\ &\quad \underline{\text{det}_k} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \underline{\text{stoch}_k} : (\iota_{i,k}, 0, 0).SC_{i,k}(\mathbf{X}) \\ E_1(\mathbf{S}) &\stackrel{def}{=} \underline{\text{init}} : (e_1, -1, k_1 S.E).E_1(\mathbf{S}) + \underline{\text{det}_1} : (e_1, -1, k_1 S.E).E_1(\mathbf{S}) + \underline{\text{stoch}_1} : (e_1, 0, 0).E_1(\mathbf{S}) \end{aligned}$$

where $J_{i,k} = 1$ if species S_i is a product in reaction k and $J_{i,k} = -1$ if species S_i is a reactant in reaction k. Furthermore $iv(\iota_{i,j}) = X_i$ and $iv(e_1) = E$. These subcomponents initially set an influence that affects variable X_i by the rate of reaction k, so the initial assumption is that reaction k is being treated deterministically. The influence is positive or negative depending on the role of S_i in the reaction. To simplify presentation, instead of using $I[\mathcal{W}]$ and defining $[\![I[\mathcal{W}]]\!]$ separately, the function $[\![I[\mathcal{W}]]\!]$ will appear explicitly in influences[8].

The subcomponents can react to two events: one which describes when the reaction becomes treated deterministically which has the same influence as the influence for the initial event and one which describes when the reaction becomes treated stochastically. In this latter case, there should be no influence in a deterministic manner and hence the influence strength and type are set to zero.

Making this concrete by considering E in the first reaction, initially the influence name e_1 is mapped to the rate function for the first reaction with a negative strength since E is consumed by the reaction. Whenever conditions are satisfied for the event $\underline{\text{stoch}_1}$ to take place, the influence is changed to a zero influence since whenever the first reaction is treated stochastically, reaction occurrences will be modelled explicitly with stochastic events that have no effect on the subcomponent. Additionally whenever the conditions are satisfied for $\underline{\text{det}_1}$, an influence (identical to the initial influence) will come into force and E will be treated deterministically in the first reaction. Note that the thresholds do not appear in the subcomponents since they are associated with the event conditions of $\underline{\text{stoch}_1}$ and $\underline{\text{det}_1}$.

[8]The original separation of influence types and their functions was to enable separation of concerns and support the idea of being able to modify the functions without modifying the subcomponents.

Next, event conditions are defined. The initial event $\underline{\text{init}}$ has a true activation condition as is required and initialises all species variables to their starting values.

$$ec(\underline{\text{init}}) = (true, \bigwedge_{i=1}^{N} X_i' = x_{i,0}) \qquad ec(\underline{\text{init}}) = (true,\ S' = S_0 \ \wedge\ E' = E_0 \ \wedge\ C' = C_0 \ \wedge\ P' = P_0)$$

The deterministic event $\underline{\text{stoch}}_k$ determines when reaction k should be treated stochastically. The reactants of this reaction need to be identified so let $R_k = \{i \mid S_i \text{ is a reactant in reaction } k\}$. The event conditions for $\underline{\text{stoch}}_k$ are defined as follows.

$$\begin{aligned} ec(\underline{\text{stoch}}_k) &= (\ f_k(\mathbf{X}) \leqslant r_k \ \vee \bigvee_{i \in R_k} X_i \leqslant h_{i,k},\ true) \\ ec(\underline{\text{stoch}}_1) &= (k_1 S.E \leqslant r_1 \ \vee\ S \leqslant h_{S,1} \ \vee\ E \leqslant h_{E,1},\ true) \qquad R_1 = \{S, E\} \end{aligned}$$

The activation condition will be true when one species or the rate is under the threshold. No resets take place upon switching, so the quantity of a species is not transformed to an integer. However, any stochastic reaction affecting that species will increment or decrement it by integer amounts. Note that other reactions may still be treating this species deterministically and hence it may also be increasing or decreasing continuously, hence if it were to be transformed to an integer, it may not remain integral.

Similarly, the deterministic event $\underline{\text{det}}_k$ determines when reaction k should be treated deterministically and has the negation of the activation condition of $\underline{\text{stoch}}_k$. However, for implementation reasons and to stop unnecessary switching between stochastic and deterministic treatment, an additional small quantity ε is added to each limit. The appropriate value for ε will be considered when the properties of the HYPE model are investigated in the next section.

$$\begin{aligned} ec(\underline{\text{det}}_k) &= (\ f_k(\mathbf{X}) > r_k + \varepsilon \ \wedge \bigwedge_{i \in R_k} X_i > h_{i,k} + \varepsilon, true\ ) \\ ec(\underline{\text{det}}_1) &= (k_1 S.E > r_1 + \varepsilon \ \wedge\ S > h_{S,1} + \varepsilon \ \wedge\ E > h_{E,1} + \varepsilon,\ true) \qquad R_1 = \{S, E\} \end{aligned}$$

When considering reaction 1 and the species E it can be seen that its quantity is compared to the its threshold value to determine whether reaction 1 should be treated stochastically or deterministically. However, it is not only the value of E that can cause $\underline{\text{stoch}}_1$ or $\underline{\text{det}}_1$ to occur, since it could also be triggered by the value of S or the value of the reaction rate. Notice that if a particular reaction becomes stochastic, only that reaction treats its species stochastically; other reactions can still treated this species deterministically. When reaction k is being treated stochastically, a stochastic event is required to represent the reaction taking place. Since a reaction can only take place if there are sufficient reactants, the rate is modified by indicator functions which return 1 only if a reactant quantity is the same as or larger than its stoichiometric coefficient in reaction k. Here products as well as reactants must be identified, hence let $Q_k = \{i \mid S_i \text{ is a product in reaction } k\}$ be the products of reaction k.

$$\begin{aligned} ec(\overline{\text{react}}_k) &= (\ f_k(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k}), \bigwedge_{i \in R_k} X_i' = X_i - \lambda_{i,k} \wedge \bigwedge_{i \in Q_k} X_i' = X_i + \lambda_{i,k}\ ) \\ ec(\overline{\text{react}}_1) &= (k_1 S.E \times I_{S,1}(1) \times I_{E,1}(1),\ S' = S - 1 \wedge E' = E - 1 \wedge C' = C + 1) \qquad R_1 = \{S, E\} \qquad Q_1 = \{C\} \end{aligned}$$

where for species i and reaction k, $\lambda_{i,k}$ is the stoichiometric coefficient. Furthermore, the indicator function is defined as follows.

$$I_{i,k}(\lambda) = \begin{cases} 1 & \text{if } X_i \geqslant \lambda \\ 0 & \text{otherwise} \end{cases} \qquad I_{E,1}(\lambda_{E,1}) = \begin{cases} 1 & \text{if } E \geqslant \lambda_{E,1} (= 1) \\ 0 & \text{otherwise} \end{cases}$$

The use of the indicator function ensures that if a reactant in a reaction (where the reactant's stoichiometric constant is κ) has a value of less than κ, will cause the rate to be zero which means that the reaction cannot occur, thus avoiding negative populations. Since the stoichiometry for E in the first reaction is 1, then reaction 1 cannot occur for any values of E in $[0, 1)$, so E can be totally consumed but cannot become negative.

Input: Bio-PEPA model with N species and K reactions of the form $S_1(x_{1,0}) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} S_N(x_{N,0})$
thresholds for reactions, $r_1, \ldots, r_k$ and thresholds of reactions in species, $h_{1,1}, \ldots, h_{N,K}$, ε
Output: Stochastic HYPE model

- Let $\mathbf{X} = \{X_1, \ldots, X_N\}$
- Let $\underline{\text{init}}$ be an event
- $reset(\underline{\text{init}})$:= "$true$"
- For i in $\{1, \ldots, N\}$
  - $reset(\underline{\text{init}})$:= $reset(\underline{\text{init}})$ + "$\wedge X_i' = x_{i,0}$"
- Let $ec(\underline{\text{init}}) = (true, reset(\underline{\text{init}}))$
- For k in $\{1, \ldots, K\}$
  - Define $ConD_k$ and $ConS_k$ as given in the text
  - Let $\underline{\text{det}_k}$, $\underline{\text{stoch}_k}$ and $\overline{\text{react}_k}$ be events
  - $act(\underline{\text{stoch}_k})$:= "$f_k \leqslant r_k$"
  - $act(\underline{\text{det}_k})$:= "$f_k > r_k + \varepsilon$"
  - $act(\overline{\text{react}_k})$:= "$f_k(X)$"
  - $reset(\overline{\text{react}_k})$:= "$true$"
  - For i in $\{1, \ldots, N\}$
    - If species S_i is a reactant in reaction k
      - Define $SC_{i,k}(\mathbf{X})$ as given in the text using influence $\iota_{i,k}$
      - $act(\underline{\text{stoch}_k})$:= $ec(\underline{\text{act}_k})$ + "$\vee S_i \leqslant h_{i,k}$"
      - $act(\underline{\text{det}_k})$:= $act(\underline{\text{det}_k})$ + "$\wedge S_i > h_{i,k} + \varepsilon$"
      - $act(\overline{\text{react}_k})$:= $act(\overline{\text{react}_k})$ + "$\times I_{i,k}(\lambda_{i,k})$"
      - $reset(\overline{\text{react}_k})$:= $reset(\overline{\text{react}_k})$ + "$\wedge X_i' := X_i - \lambda_{i,k}$"
    - If species S_i is a product in reaction k
      - Define $SC_{i,k}(\mathbf{X})$ as given in the text using influence $\iota_{i,k}$
      - $reset(\overline{\text{react}_k})$:= $reset(\overline{\text{react}_k})$ + "$\wedge X_i' := X_i + \lambda_{i,k}$"
    - Let $iv(\iota_{i,k}) = X_i$
  - Define $RC_k(\mathbf{X})$ as given in the text
  - Let $ec(\underline{\text{stoch}_k}) = (act(\underline{\text{stoch}_k}), true)$
  - Let $ec(\underline{\text{det}_k}) = (act(\underline{\text{det}_k}), true)$
  - Let $ec(\overline{\text{react}_k}) = (act(\overline{\text{react}_k}), reset(\overline{\text{react}_k}))$
- Define BP as given in the text

Figure 5: Algorithm to transform a Bio-PEPA model to a HYPE model

Other approaches for dealing with negative populations in the context of tau-leaping have been suggested, including rounded to the nearest integer [68, 19]. However, rounding can increase or decrease the mass of the system, and the approach taken with HYPE seems preferable. Another alternative is to make adjustments at the switch that ensures a balance for that reaction, so that any increase in the reactant species is matched by a decrease in the product species for that reaction (and vice versa). But as mentioned above, the quantity of a species may still vary continuously if only one of the reactions in which it is involved is being treated stochastically, so changing to integer amounts is an unnecessary overhead.

The subcomponents and events have been defined for the HYPE model, and only the controller remains. The individual subcontrollers determine the ordering of events for each reaction and have the following definition. They ensure the alternation of stochastic and deterministic treatment of reactions, and when a reaction is being treated stochastically, its controller allows the reaction to occur.

$$ConD_k \overset{def}{=} \underline{\text{stoch}_k}.ConS_k \qquad ConS_k \overset{def}{=} \overline{\text{react}_k}.ConS_k + \underline{\text{det}_k}.ConD_k$$

Subcomponents are grouped into components: $RC_k(\mathbf{X}) \overset{def}{=} \underset{*}{\bowtie}_{i \in R_k \cup Q_k} SC_{i,k}(\mathbf{X})$. The controlled system is defined as

$$BP \overset{def}{=} (\, RC_1(\mathbf{X}) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} RC_k(\mathbf{X}) \,) \underset{*}{\bowtie} \underline{\text{init}}.(\, ConD_1 \parallel \ldots \parallel ConD_k \,)$$

$$\mathbf{S} = (S, E, C, P) \qquad iv(s_1) = iv(s_2) = S \qquad iv(e_1) = iv(e_2) = iv(e_3) = E$$
$$iv(p_3) = P \qquad iv(c_1) = iv(c_2) = iv(c_3) = C$$

$$S_1(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (s_1, -1, k_1 S.E).S_1(\mathbf{S}) + \underline{\text{stoch}_1} : (s_1, 0, 0).S_1(\mathbf{S}) + \underline{\text{det}_1} : (s_1, -1, k_1 S.E).S_1(\mathbf{S})$$
$$E_1(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (e_1, -1, k_1 S.E).E_1(\mathbf{S}) + \underline{\text{stoch}_1} : (e_1, 0, 0).E_1(\mathbf{S}) + \underline{\text{det}_1} : (e_1, -1, k_1 S.E).E_1(\mathbf{S})$$
$$C_1(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (c_1, 1, k_1 S.E).E_1(\mathbf{S}) + \underline{\text{stoch}_1} : (c_1, 0, 0).C_1(\mathbf{S}) + \underline{\text{det}_1} : (c_1, 1, k_1 S.E).C_1(\mathbf{S})$$

$$C_2(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (c_2, -1, k_2 C).E_2(\mathbf{S}) + \underline{\text{stoch}_2} : (e_2, 0, 0).C_2(\mathbf{S}) + \underline{\text{det}_2} : (e_2, -1, k_2 C).C_2(\mathbf{S})$$
$$S_2(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (s_2, 1, k_2 C).S_2(\mathbf{S}) + \underline{\text{stoch}_2} : (s_2, 0, 0).S_2(\mathbf{S}) + \underline{\text{det}_2} : (s_2, 1, k_2 C).S_2(\mathbf{S})$$
$$E_2(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (e_2, 1, k_2 C).E_2(\mathbf{S}) + \underline{\text{stoch}_2} : (e_2, 0, 0).E_2(\mathbf{S}) + \underline{\text{det}_2} : (e_1, 1, k_2 C).E_2(\mathbf{S})$$

$$C_3(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (c_3, -1, k_3 C).E_3(\mathbf{S}) + \underline{\text{stoch}_3} : (e_3, 0, 0).C_3(\mathbf{S}) + \underline{\text{det}_3} : (e_3, -1, k_3 C).C_3(\mathbf{S})$$
$$P_3(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (p_3, 1, k_3 C).P_3(\mathbf{S}) + \underline{\text{stoch}_3} : (p_3, 0, 0).S_3(\mathbf{S}) + \underline{\text{det}_3} : (p_3, 1, k_3 C).P_3(\mathbf{S})$$
$$E_3(\mathbf{S}) \stackrel{def}{=} \underline{\text{init}} : (e_3, 1, k_3 C).E_3(\mathbf{S}) + \underline{\text{stoch}_3} : (e_3, 0, 0).E_3(\mathbf{S}) + \underline{\text{det}_3} : (e_1, 1, k_3 C).E_3(\mathbf{S})$$

$$ec(\underline{\text{init}}) = (true,\ S' = S_0 \wedge E' = E_0 \wedge C' = C_0 \wedge P' = P_0)$$

$$ec(\underline{\text{stoch}_1}) = (k_1 S.E \leqslant r_1 \vee S \leqslant h_{S,1} \vee E \leqslant h_{E,1},\ true)$$
$$ec(\underline{\text{stoch}_2}) = (k_2 C \leqslant r_2 \vee C \leqslant h_{C,2},\ true) \qquad ec(\underline{\text{stoch}_3}) = (k_3 C \leqslant r_3 \vee C \leqslant h_{C,3},\ true)$$

$$ec(\underline{\text{det}_1}) = (k_1 S.E > r_1 + \varepsilon \wedge S > h_{S,1} + \varepsilon \wedge E > h_{E,1} + \varepsilon,\ true)$$
$$ec(\underline{\text{det}_2}) = (k_2 C > r_2 + \varepsilon \wedge C > h_{C,2} + \varepsilon,\ true) \quad ec(\underline{\text{det}_3}) = (k_3 C > r_3 + \varepsilon \wedge C > h_{C,3} + \varepsilon,\ true)$$

$$ec(\overline{\text{react}_1}) = (k_1 S.E \times I_{S,1}(1) \times I_{E,1}(1),\ S' = S - 1 \wedge E' = E - 1 \wedge C' = C + 1)$$
$$ec(\overline{\text{react}_2}) = (k_2 C \times I_{C,2}(1),\ S' = S + 1 \wedge E' = E + 1 \wedge C' = C - 1)$$
$$ec(\overline{\text{react}_3}) = (k_3 C \times I_{C,3}(1),\ P' = P + 1 \wedge E' = E + 1 \wedge C' = C - 1)$$

$$ConD_k \stackrel{def}{=} \underline{\text{stoch}_k}.ConS_k \qquad ConS_k \stackrel{def}{=} \overline{\text{react}_k}.ConS_k + \underline{\text{det}_1}.ConD_k \quad k \in \{1, 2, 3\}$$

$$R_1(\mathbf{S}) \stackrel{def}{=} S_1(\mathbf{S}) \underset{*}{\bowtie} E_1(\mathbf{S}) \underset{*}{\bowtie} C_1(\mathbf{S}) \qquad R_2(\mathbf{S}) \stackrel{def}{=} S_2(\mathbf{S}) \underset{*}{\bowtie} E_2(\mathbf{S}) \underset{*}{\bowtie} C_2(\mathbf{S})$$
$$R_3(\mathbf{S}) \stackrel{def}{=} P_3(\mathbf{S}) \underset{*}{\bowtie} E_3(\mathbf{S}) \underset{*}{\bowtie} C_3(\mathbf{S})$$

$$EK \stackrel{def}{=} (R_1(\mathbf{S}) \underset{*}{\bowtie} R_2(\mathbf{S}) \underset{*}{\bowtie} R_3(\mathbf{S})) \underset{*}{\bowtie} \underline{\text{init}}.(ConD_1 \parallel ConD_2 \parallel ConD_3)$$

Figure 6: Stochastic HYPE model for enzyme kinetics example

where $\parallel$ is shorthand for $\underset{\emptyset}{\bowtie}$. The process of transforming a Bio-PEPA model to a HYPE model can be described algorithmically and this process is described using pseudocode in Figure 5. The formal model is described in Appendix A. As described above, each species can be affected by continuous flows and stochastic reactions simultaneously. All reactions start as deterministic, and will immediately switch to stochastic treatment if the thresholds warrant it.

5.1. Example revisited: enzyme kinetics

The transformation of the enzyme example is given in Figure 6. It is not possible to give the full TDSHA for this model; however, two modes are now described. The first mode that is reached after the $\underline{\text{init}}$ event has the system of ODEs given in Section 3.2 since initially all reactions are treated deterministically. It has three outgoing instantaneous transitions for the events $\underline{\text{stoch}_1}$, $\underline{\text{stoch}_2}$ and $\underline{\text{stoch}_3}$, and three incoming transitions for the events $\underline{\text{det}_1}$, $\underline{\text{det}_2}$ and $\underline{\text{det}_3}$. Assuming that after the passing of some time, $\underline{\text{stoch}_1}$ is triggered. The new mode has the following ODEs.

$$dS/dt = +k_2 C \qquad dE/dt = +k_2 C + k_3 C \qquad dP/dt = +k_3 C \qquad dC/dt = -k_2 C - k_3 C$$

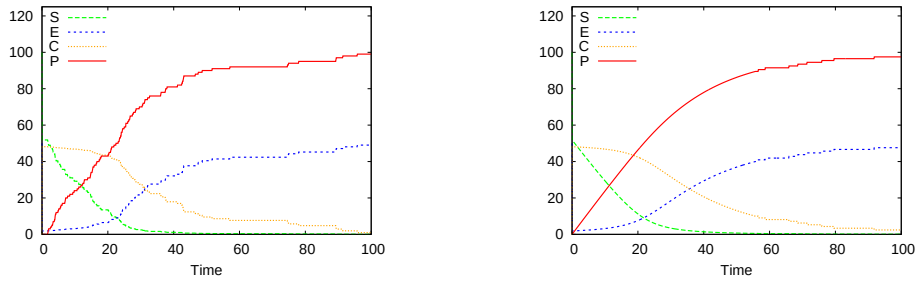

Figure 7: Simulation of the HYPE *EK* model using reaction rate thresholds of 50 for all reactions (left) and simulation of the same model using a threshold of 10 for C in reaction 3 (right) with parameters and initial values as in Figure 3 and $\varepsilon = 1$

It has a stochastic transition for $\overline{\text{react}}_1$ which loops back to the new mode. Additionally, there are outgoing transitions for events $\underline{\text{stoch}}_2$ and $\underline{\text{stoch}}_3$ and an outgoing event $\underline{\text{det}}_1$ whose target is the first mode. There are also incoming transitions for events $\underline{\text{det}}_2$ and $\underline{\text{det}}_3$. In the first mode, everything is treated deterministically, and in the second, the first reaction is treated stochastically, leading to the species S, E and C being modified both continuously and in unit steps, whereas P is only modified continuously since it is not involved in the first reaction.

As mentioned in Section 3.1, this example typically has a combination of fast rates and slow rates, and this is the case for this model since k_1 and k_2 are much larger than k_3. This suggests that reactions k_1 and k_2 can be modelled deterministically and k_3 stochastically and this is achieved by setting the reaction thresholds to $r_1 = r_2 = r_3 = 50$, and all other thresholds to negative values. The output of a single simulation run using these thresholds are given in Figure 7. All simulations of the HYPE model were generated by the stochastic hybrid simulator described in [8].

Another possibility is to set a threshold that is triggered by low molecule counts. In the right-hand graph in Figure 7, all thresholds are set to negative values except that for species C in reaction 3 which is set to 10. It can be seen that as soon as C reaches 10 molecules, reaction 3 is treated stochastically which also can also been seen in the lines for E and P since they are both involved as products in reaction 3. S is not affected directly.

The next section considers the properties of the generated HYPE model, after which a case study is presented.

5.2. Properties of the HYPE model

The following proposition considers the size of the HYPE model obtained from a Bio-PEPA model.

Proposition 3. *Given a well-defined Bio-PEPA model with N species and K reactions, the HYPE model obtained from this Bio-PEPA model has*

- *N variables*
- *at most $N \times K$ subcomponents with 3 terms*
- *at most $N \times K$ influences*
- *$2K + 1$ instantaneous events*
- *K stochastic events*
- *K components*
- *$2K$ controllers with at most 2 terms*

Since the size of a Bio-PEPA model with N species and K reactions is N species initial values, K reaction functions and at most $N \times K$ reaction definitions, the size of the HYPE model is at most 6 times larger than the Bio-PEPA model, and hence there is no exponential (or even polynomial) blow-up in the mapping to HYPE.

This generic HYPE model satisfies all of the requirements for being well-defined except for the requirement that $\mathsf{ev}(Con) \subseteq \mathsf{ev}(Reactions)$.

Lemma 1. *Given a well-defined Bio-PEPA system $\mathcal{P}$, its HYPE model is well-defined with the addition of the subcomponent $W(\mathbf{X}) \stackrel{\text{def}}{=} \underline{\text{init}} : (w, 0, 0).W(\mathbf{X}) + \sum_{k=1}^{K} \overline{\text{react}}_k : (w, 0, 0).W(\mathbf{X})$ with a new variable W and $iv(w) = W$.*

Proof. First note that all subcomponents have the required form and specifically, each influence is unique to its subcomponent. The operator $\bowtie_*$ is used to ensure synchronisation on all events between subcomponents and also between the uncontrolled system and the controller. The controlled system has the correct form since the controllers are prefixed by $\underline{\text{init}}$. The additional subcomponent is required to ensure that every event in the controllers appears in the uncontrolled system but it has no effect on the behaviour of the system.

Lemma 2. *Given a well-defined Bio-PEPA system $\mathcal{P}$, its HYPE model is well-behaved.*

Proof. To determine if the model is well-behaved, its controllers and event conditions must be inspected to see if it is possible for a infinite sequence of instantaneous events to occur at a single time point (instantaneous Zeno behaviour).

Each $ConD_k$ can be shown to be well-behaved using ideas similar to those in Proposition 1. Each controller can be viewed as providing a sequence of $\underline{\text{stoch}}_k$ events interleaved with $\overline{\text{react}}_k$ events and $\underline{\text{det}}_k$ events. Since $\overline{\text{react}}_k$ is a stochastic event, its occurrence will definitely end any sequence of instantaneous actions. Is it possible for there to be an infinite sequence of $\underline{\text{stoch}}_k$ and $\underline{\text{det}}_k$ at a time instant without a $\overline{\text{react}}_k$ event occurring? Note first that neither $\underline{\text{stoch}}_k$ nor $\underline{\text{det}}_k$ change the values of any variables as they both have *true* resets. Hence, the activation conditions of these two events do not overlap, and this means that time must elapse between them (in either order) because the values of variables must change. Clearly, any value that activates $\underline{\text{stoch}}_k$ will not activate $\underline{\text{det}}_k$, and *vice versa*. Hence each $ConD_k$ is well-behaved.

Next consider a pair of controllers $ConD_i$ and $ConD_j$. Their events are disjoint, and no event can activate another, because all resets are the identity, hence if an event is active after another event, it must have been active before that event as well. Hence by repeated use of Proposition 2, the whole controller is well-behaved.

This result covers only instantaneous behaviour at a single time point. In Section 5.3, another undesirable behaviour that can occur over a period of time and make simulation much slower will be discussed, and a calculation will be given for ε to reduce the chance of this behaviour occurring.

Next, the two extremes of the HYPE model are considered and are proved to be the same as the original Bio-PEPA model, considered either deterministically or stochastically. A definition about reaction rate functions is required first.

Definition 2. Let f be the rate function for a reaction with reactants $A_1, \ldots, A_n$. The function $f(A_1, \ldots, A_n, \ldots)$ is *well-behaved* if whenever $A_i = 0$ for some $i \in \{1, \ldots, n\}$ then $f(A_1, \ldots, A_n, \ldots) = 0$

The first result considers the situation where all thresholds are below zero.

Theorem 1. *If for each species S_i, $h_{i,k} \leq 0$; for each reaction k, $r_k \leq 0$; and every rate function is well-behaved, then the HYPE model will provide a purely deterministic trace, and it is the same deterministic trace as generated by the original Bio-PEPA model.*

Proof. First, note that if every threshold is less than zero, then none of the events $\underline{\text{stoch}}_k$ can occur, since species quantities are always non-negative. Species quantities are initially non-negative and rate functions are well-behaved hence a rate will become zero as soon as a reactant becomes zero and no further decrease in that reactant is possible.

Considering the form of each controller $ConD_k$, if event $\underline{\text{stoch}}_k$ cannot happen, then neither can $\underline{\text{det}}_k$ nor $\overline{\text{react}}_k$. The HYPE model can then be modified to one where both the $\underline{\text{stoch}}_k$ and $\underline{\text{det}}_k$ prefixes can be removed from each $SC_{i,k}$ leaving

$$SC'_{i,k}(\mathbf{X}) \quad = \quad \underline{\text{init}} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC'_{i,k}(\mathbf{X})$$

Since no events except $\underline{\text{init}}$ can happen, the controller can be reduced to 0 giving the system

$$BP' \quad \stackrel{def}{=} \quad (\ RC'_1(\mathbf{X}) \bowtie_* \ \ldots \ \bowtie_* RC'_k(\mathbf{X})\ ) \bowtie_* \underline{\text{init}}.0.$$

The TDSHA of the new HYPE model BP' will have the same behaviour as the TDSHA of BP under the conditions of negative thresholds since modification has only removed events that cannot happen. The labelled multitransition system of BP' will have exactly one configuration $\langle (\ RC'_1(\mathbf{X}) \bowtie_* \ \ldots \ \bowtie_* RC'_k(\mathbf{X})\ ) \bowtie_* 0, \sigma \rangle$ after $\underline{\text{init}}$, and when this

is transformed to a TDSHA mode, it will have continuous behaviour defined by the following ODEs, one for each species S_i,

$$\begin{aligned}\frac{dX_i}{dt} &= \sum\{r \cdot f(\mathbf{X}) \mid iv(\iota) = X_i \text{ and } \sigma(\iota) = (r, f(\mathbf{X}))\} \\ &= \sum\{-\lambda_{i,k} \times f_k(\mathbf{X})) \mid S_i \text{ is a reactant in reaction } k\} + \sum\{\ \lambda_{i,k} \times f_k(\mathbf{X})) \mid S_i \text{ is a product in reaction } k\}\end{aligned}$$

since each $\iota_{i,k}$ is associated with variable X_i and in state σ, $\sigma(\iota_{i,k}) = (\lambda_{i,k} \times J_{i,k}, f_k(\mathbf{X}))$. These ODEs are exactly those obtained from the original Bio-PEPA model of the form

$$\frac{d\mathbf{X}}{dt} = \mathbf{D} \times \mathbf{v}$$

where $\mathbf{D} = \{d_{i,k}\}$ is the $N \times K$ stoichiometry matrix and $\mathbf{v}$ is a K-length vector containing the kinetic laws for each reaction with each element of the form $f_k(\mathbf{X})$. Since the HYPE model has no instantaneous behaviour (after the occurrence of $\underline{\text{init}}$) and no stochastic behaviour, its TDSHA will only exhibit continuous behaviour and this continuous behaviour is defined by the same system of ODEs that defines the continuous behaviour of the Bio-PEPA system, hence the two models will generate the same deterministic trace. □

Next, the case of infinite thresholds is considered. This ensures that the model will be purely stochastic after $\underline{\text{init}}$ and each of the $\underline{\text{stoch}}_k$ events have occurred at time zero.

Theorem 2. *If for each species S_i, its thresholds $h_{i,k}$ are infinite; for each reaction k, its threshold are infinite; and every rate function is well-behaved, then the CTMC defined by the HYPE model is equivalent to the CTMC defined by considering the original Bio-PEPA model stochastically.*

Proof. The CTMC obtained from the Bio-PEPA model is an infinite CTMC with states of the form $(x_1, \ldots, x_n)$ where each x_i represents an integral quantity of species S_i (assuming integral initial values).

A transition t_k representing the occurrence of reaction k is defined from the state $(x_1, \ldots, x_n)$ to the state $(x_1 + (\lambda_{1,k} \times J_{1,k}), \ldots, x_n + (\lambda_{N,k} \times J_{N,k}))$ where any $\lambda_{i,k}$ and $J_{i,k}$ not already defined have value zero.

The transition t_k can only occur if there is sufficient number of the reactants in the state $(x_1, \ldots, x_n)$ and it has rate $f_k(\mathbf{X})$. Using the definition of R_k for the reactants in reaction k and the comparison function $I_{i,k}(\lambda_{i,k})$ as defined above together with the fact that f_k is well-behaved then

$$f_k(\mathbf{X}) = f_k(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k})$$

whenever there are sufficient reactants.

The labelled multitransition system of the HYPE model obtained from the Bio-PEPA model has a number of configurations but the main configuration of interest is the one that is reached after all $\underline{\text{stoch}}_k$ have occurred once in any order, and it has the form $q = \langle RC_1(\mathbf{X}) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} RC_k(\mathbf{X})) \underset{*}{\bowtie} (\ \overline{ConS_1} \parallel \ldots \parallel \overline{ConS_k}, \sigma\rangle$. The only events that are possible in this mode in the TDSHA are the stochastic events $\overline{\text{react}}_k$ because the $\underline{\text{det}}_k$ events will never be activated due to infinite thresholds. Also note that in the behaviour of the TDSHA, the instantaneous events $\underline{\text{stoch}}_k$ have all happened immediately after $\underline{\text{init}}$ since their activation conditions (being under the thresholds) are immediately satisfied. Hence at time zero, the system is immediately in mode q. Hence the remaining behaviour (immediately after time zero) is defined by the occurrence of the stochastic events $\underline{\text{react}}_k$. In the TDSHA, these give rise to stochastic transitions from mode q to itself. Since initial values of species quantities are integral, and any changes to these values via the resets of each $\underline{\text{react}}_k$ are determined by stoichiometric coefficients (which are integral), all behaviour of the TDSHA only involves states of the form $(q, (x_1, \ldots, x_n))$ where each x_i is integral (and non-negative). Since the stochastic transitions use stoichiometric coefficients in their resets, the transitions are the same transitions as in the CTMC generated by the Bio-PEPA model. Finally the transition rate for $\overline{\text{react}}_k$ is $f_k(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k})$ ensuring that transitions only occur when there are sufficient reactants and matching the rates in the CTMC generated by the Bio-PEPA model. □

This theorem assumes that there are no upper bounds on the quantity of each species and hence the CTMC can be infinite. In the case that maximum quantities are imposed then like with Bio-PEPA it would be possible to map the HYPE model (in its purely stochastic form) to PRISM [53] and then model checking can be performed.

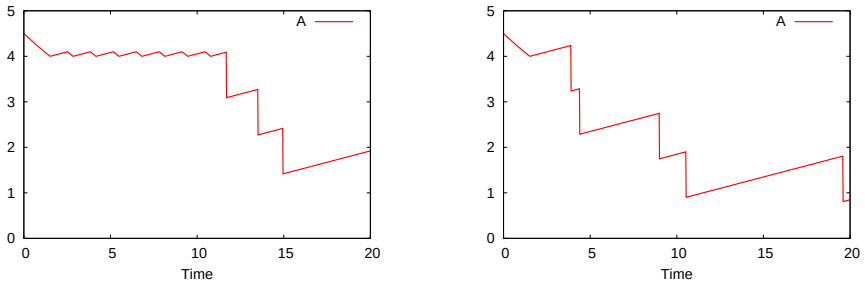

Figure 8: Hybrid simulation showing how switching can occur when ε is too small (left) and how a larger value can mitigate against this (right) $k_1 = 0.1$, $k_2 = 0.1$, $h_{A,2} = 4$ and $\varepsilon = 0.1$ (left) $\varepsilon = 0.4$ (right)

5.3. Choice of ε for simulation

The following simple example illustrates the technical necessity of the use of ε in event conditions which switch from stochastic treatment of a reaction to deterministic treatment of that reaction. Consider a single species with a reaction that creates the species and a reaction that degrades it. In Bio-PEPA, this can be written as $A \stackrel{def}{=} (r_1, 1)\uparrow A + (r_2, 1)\downarrow A$. Reaction r_1 has a constant rate k_1 and the rate for reaction r_2 is based on mass action with definition $k_2 A$.

Considering the HYPE model obtained from this Bio-PEPA model, let $h_{A,2}$ be the threshold quantity of A in reaction r_2 such that if $A \leq h_{A,2}$, reaction r_2 is treated stochastically. First, consider the case when $\varepsilon = 0$. In this case, the treatment of r_2 will switch back to deterministic when $A > h_{A,2}$. Let the rate of r_2 be larger than the rate of r_1. This implies that $dA/dt < 0$ and the amount of A is decreasing. Assume that in the simulation of the HYPE model, A has decreased to $h_{A,2}$. This will be detected either when the quantity of A is at or below $h_{A,2}$ (because of the imprecise nature of numerical simulation, it may not detect it exactly at equality) and switch treatment of r_2 to stochastic. Since the deterministic contribution of r_2 is now removed, A will be increasing at rate k_1 and $dA/dt > 0$. The simulation will proceed, and even if the next step is small, it is very likely that A will be over the threshold, leading to deterministic treatment of r_2. Hence, the quantity of A will start decreasing and the simulation will switch to a stochastic treatment of r_2 and this cycle will continue *ad infinitum*. Introducing ε effectively splits the single threshold into two distinct thresholds and through the difference in values, prevents this unnecessary cycling.

Since the mapping to HYPE from Bio-PEPA is motivated by efficiency, the exclusion of unnecessary events in the simulation of the HYPE model is important. Even with a non-zero ε, the switching between stochastic and deterministic treatment of the reaction r_2 without the occurrence of reaction r_2 events can occur until the value of A is low enough that it becomes very unlikely that it will increase to above the threshold before reaction r_2 occurs again. The left-hand graph in Figure 8 illustrates how this can occur for species A as defined above. The simulation shows repeated switching. It is only after the first r_2 reaction at approximately time unit 19 that the switching stops.

This raises the question of what value to choose for ε. For the current example, one would like to bound the probability that threshold is passed before the next r_2 reaction. In other words, one would like to fix the probability that the quantity of A increases enough to be over the threshold before an r_2 reaction occurs, as this will bound the probability of unnecessary switching. Since A is increasing at rate k_1 when only r_1 is treated deterministically, ε should be a multiple of k_1 where the multiplier is the amount of time during which there is a p probability of reaction r_2 occurring.

The function $F_\lambda(x) = 1 - e^{-\lambda x}$ defines the cumulative distribution of an exponential distribution with parameter λ. Its inverse $F_\lambda^{-1}(p)$ takes a probability as its argument and returns the time at which there is probability p that the event with duration specified by λ will have occurred. This function can be used to estimate a value for ε.

Let $\varepsilon = F^{-1}_{k_2 A'}(p) \times k_1$ where p is the desired likelihood that the duration of the next reaction r_2 is longer than the number of time steps it will take A' (the current quantity of species A) to increase to $A' + \varepsilon$. In the left-hand graph in Figure 8, $\varepsilon = 0.1$ corresponds with a 0.33 probability that the reaction will occur before the amount of A has increased by 0.1, and reached 4.1. By contrast, in the right-hand graph, $\varepsilon = 0.4$ which corresponds to a probability of 0.8. For the particular simulation given in this graph, the value of A has increased significantly after the threshold for changing to stochastic treatment but is still less than the threshold when the reaction r_2 occurs.

	Reaction						Rate function
1				$\longrightarrow$	G_α		k_1
2		G_α		$\longrightarrow$	$2G_\alpha$		$k_2 G_\alpha$
3	PLC^* +	G_α		$\longrightarrow$		PLC	$k_3 G_\alpha PLC^*/(K_4 + G_\alpha)$
4	Ca +	G_α		$\longrightarrow$		Ca	$k_5 G_\alpha Ca/(K_6 + G_\alpha)$
5		G_α		$\longrightarrow$	G_α +	PLC^*	$k_7 G_\alpha$
6	PLC^*			$\longrightarrow$			$k_8 PLC^*/(K_9 + PLC^*)$
7		G_α		$\longrightarrow$	G_α +	Ca	$k_{10} G_\alpha$
8	Ca			$\longrightarrow$			$k_{11} Ca/(K_{12} + Ca)$

Table 1: Reactions and reaction rates for the calcium oscillations in hepatocytes [46]

To consider this more generally, the unnecessary switching could take place in any system of reactions where there is a species S_i which is a reactant in reaction r_k where $dS_i/dt < 0$, r_k is being treated deterministically and S_i is slightly larger than $h_{i,k}$. Moreover, after switching to stochastic treatment of r_k, it must be the case that $dS_i/dt > 0$ (to obtain unnecessary switching). Let this rate of change be δ then an estimate for ε is $F^{-1}_{f_{i,k}(\mathbf{X})}(p) \times \delta$ where p is the required probability. This is necessarily an estimate, since $f_{i,k}$ is changing over time and dependent on all reactant species in reaction k.

Because $F^{-1}_{f_{i,k}(\mathbf{X})}(p)$ is calculated for $\mathbf{X}$ at a single point in time, rather than for a distribution of values over time, the smaller the change in $f_{i,k}$, the more accurate the estimate is likely to be. This can occur in two situations. In the first, a better estimate will be obtained when the ODEs for the reactant species are changing slowly, as this means the value of $f_{i,k}$ will not change fast. Alternatively, a better estimate will be obtained if $f_{i,k}$ is fast and hence there is little opportunity for the values of the reactant species to change.

In practice, it is unlikely that one can straightforwardly determine from the model before simulation whether unnecessary switching does occur, and an inspection of the simulation trace may be necessary to identify it[9]. In many models, it may not occur at all, and here, a small positive value can be used for ε. In other models, it may occur frequently and hence the above estimation can be used to determine a value. In some models, it may occur for more than one reaction, and different values of ε may be needed for different reactions, or even different species within different reactions.

6. Case study

This section considers a model of calcium oscillation in liver cells [51] as presented in [46]. This model is presented in Table 1. There are three species and eight reactions. The first reaction represents the activation of the α subunit of the G protein (G_α). The second reaction is the self-activation of G_α and the value of its parameter k_2 is key in determining whether the system oscillates. Reactions 3 and 4 represent the degradation of G_α enzymatically by the activated form of phospholipase C (PLC^*) and cytosolic calcium ions (Ca) respectively. These two species are catalysed by G_α (reactions 5 and 7) and also have simple degradation reactions (reactions 6 and 8). There are feedback loops in this model, and hence oscillations can occur for the right parameter ranges.

With the parameters presented in [51, 46], this model shows oscillatory behaviour in deterministic analysis and stochastic analysis. In Section 2, quasicycles were discussed as a feature of models under certain parameters where stochastic analysis reveals oscillatory behaviour but no limit cycle was present (no oscillatory behaviour under deterministic simulation). Although the model with the given parameters does not fit this description, it can be modified to demonstrate quasicycles. Harris *et al* [46] note that k_2 is the parameter that relates to the oscillatory behaviour. Depending on the number of molecules, the modification of this parameter can result in quasicycles. The question which is then raised is how a hybrid interpretation can be used where most reactions are modelled deterministically, some are modelled stochastically (when appropriate) and the oscillatory behaviour is recovered.

[9]The simulator used for the models in this paper has an option to display each event as it occurs and unnecessary switching can easily be detected in this display.

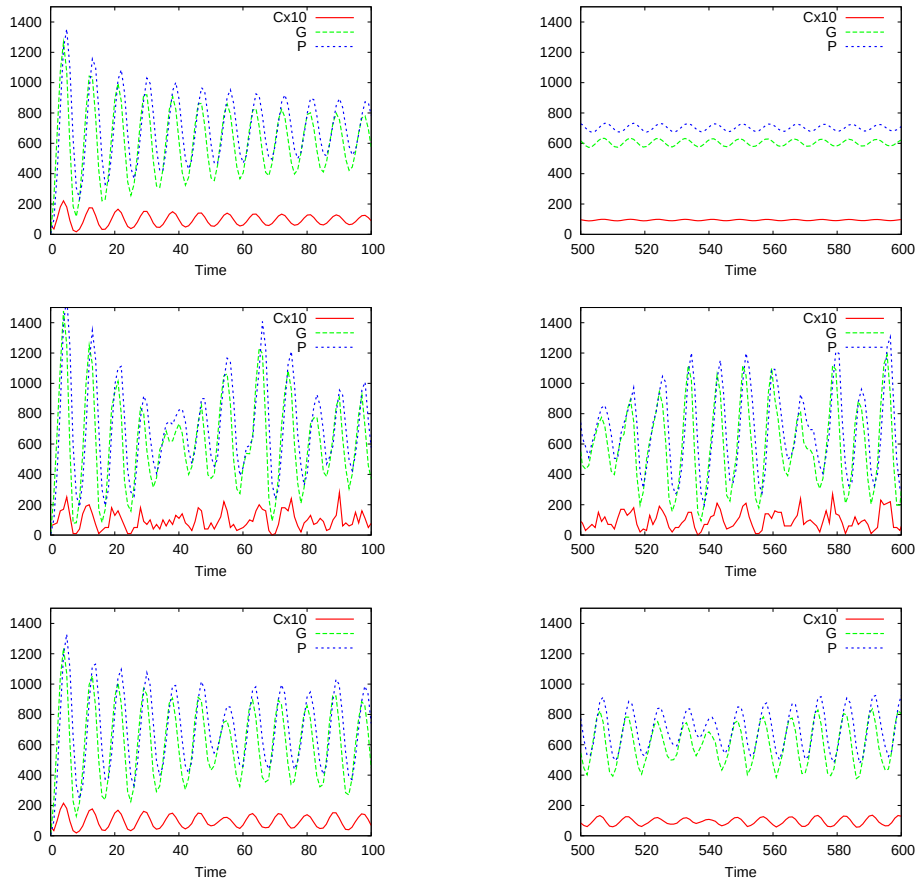

Figure 9: Simulation of the Bio-PEPA calcium model showing deterministic simulation with damping oscillations (top row), stochastic simulation with persistent oscillations (middle row) and simulation of the HYPE calcium model with persistent oscillations with $\varepsilon = 0.2$ and all thresholds at -1 except for $h_{G,2} = 625$ (bottom row) all graphs use the obvious abbreviations for the three species and parameters as given in [46] except $k_2 = 2.85 \times 0.46$

The model was first expressed in Bio-PEPA and then translated to a HYPE model using the mapping proposed in this paper. For the original parameters, both stochastic and deterministic simulations agreed with those of the earlier papers. The parameter k_2 was reduced by multiplying it by 0.46 to obtain the situation where the stochastic simulation (of the Bio-PEPA model) showed persistent oscillations and the deterministic simulation (of the Bio-PEPA model) showed damped oscillations as illustrated in the top row of Figure 9 (the value of Ca has been multiplied by 10 so that it is more visible in the graph). The HYPE model was also simulated with all thresholds set to -1 so that no reactions would be treated stochastically and demonstrated the same damped oscillation as the deterministic simulation of the Bio-PEPA model (not shown), thus illustrating the validity of the HYPE model.

Various investigations were carried out to see what thresholds to choose and the graph presented in Figure 9 show the case where $h_{G,2} = 625$. The bottom row of Figure 9 show how the oscillations are now persistent through treating certain reactions stochastically. In effect, sufficient noise is introduced to remove the damping affect, although the oscillations do not have the same amplitude as those of the fully stochastic model. The value $h_{G,2} = 625$ shows very persistent oscillations, explored by runs of 10000 time units. By contrast, a lower threshold of 600 shows initial oscillations but they disappear suddenly (without damping) as shown in Figure 10. Experiments with adding thresholds for other species in other reactions did not change this behaviour, except that adding a threshold of $h_{C,4} = 20$ (the amount of calcium in the reaction that degrades G_α) appears to ensure that oscillations will appear again although the graph becomes very flat in places (Figure 10). A value of 0.2 was used for ε and investigating the traces of events,

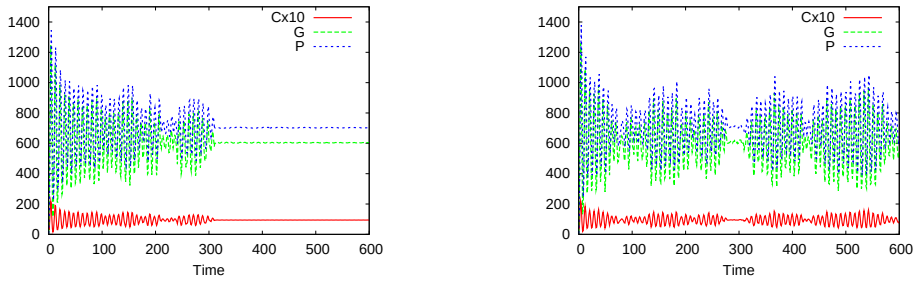

Figure 10: Simulation of the HYPE calcium model with $\varepsilon = 0.2$ and all thresholds at -1 except for $h_{G,2} = 600$ (left) with $\varepsilon = 0.2$ and all thresholds at -1 except for $h_{G,2} = 600$ and $h_{C,4} = 20$ (right) using the obvious abbreviations for the three species and parameters as given in [46] except $k_2 = 2.85 \times 0.46$

there was no sign of unnecessary cycling and hence a low value is appropriate for this model.

Since the software for simulating Bio-PEPA and HYPE are two different pieces of software it is not possible to compare runtimes directly. Considering just the HYPE simulation with a simulation running for 1000 time units, and estimating runtime of a fully stochastic version of model as more than 12 hours, the runtime of the hybrid simulation of about 1 minute shows a significant reduction over the fully stochastic model. Another approach to this analysis is to consider a system of K reactions and assume they have similar rates, the cost of deterministic simulation is negligible and that are approximately M events in a fully stochastic simulation of a fixed length. If only one of these reactions is treated stochastically for the whole of the simulation then the number of stochastic events is reduced to M/K. Furthermore, if that reaction is treated stochastically only for a fraction of the simulation then there is further reduction.

7. Discussion

This paper has introduced a method of translating Bio-PEPA models to HYPE models to allow for stochastic simulation with the aim of accessing the advantages of deterministic simulation and stochastic simulation together and removing their disadvantages. As the case study in the previous section shows, it is possible to use stochastic treatment of reactions which is determined by thresholds to obtain behaviour that might be lost in a purely deterministic simulation without requiring full stochastic treatment.

However, it is also clear that this method when used practically requires tuning as the discussion about the choice of ε demonstrates. Additionally, understanding of the model is important. In the case study, existing research provided the understanding of the roles of each reaction and the crucial role of one parameter and reaction. These pointed to what could be modified, both to understand how the oscillations came about and which reactions might be appropriate for thresholds. An area for further work is how to use the model structure to determine where intervention is appropriate. An incorrect choice of reaction and species for a threshold can lead to a very slow simulation with no benefits. Considering the case study above, setting a threshold for Ca in reaction 8 (which is a fast reaction) is pointless. Another question raised by the case study was whether thresholds are the only technique that can be used when determining whether to switch to stochastic behaviour.

It has been proved that in the two extreme cases, the simulation of the HYPE model is the same as the simulation of the Bio-PEPA model. In the deterministic case the same ODEs are obtained and in the stochastic case, the same CTMCs are obtained. This raises the question of what happens for intermediate cases. To determine if both extreme cases show the same behaviour then any intermediate case will also show that behaviour is an issue for further research. This problem can be tackled both theoretically and experimentally.

The next two sections now look at two extensions to Bio-PEPA and how these new features added to Bio-PEPA can be included in the HYPE model generated. Both of these show that the stochastic hybrid formalism offers the relevant behaviour types and furthermore that the HYPE language provides constructs to conveniently express these features.

8. Bio-PEPA with events

In [26] Ciocchetta extended Bio-PEPA with events, where the semantics were expressed as hybrid automata. This can also be achieved through the mapping to HYPE. In the original proposal, syntactically, a set of events is added to the tuple of a Bio-PEPA system [26]. This was done to provide Bio-PEPA with same functionality as SMBL which allowed for delays in the form now specified.

Each event has the form $(id, trigger, event\_assignment, delay)$ where id identifies the event.

trigger: This is a boolean expression defined over comparisons of species quantities, time and constants, and determines when the event happens.

event_assignment: This is the result of the event occurring and involves assignment of new values to any one of species quantity, constant, rate function, time or volume. Let assignments involving species quantities be denoted *species_assignment*.

delay: This enables the effect of an event to take place a fixed time after the event has been triggered. The use of delays in a different context is discussed further in the next section where a proposal for mapping Bio-PEPA with delays (Bio-PEPAd) [24] to HYPE is outlined. As described in [26], an event with a delay can be split into two events with zero delay. The first event has the trigger of the original event and its assignment records the current time. The trigger of the second compares the sum of the recorded time and the delay with the current time, and the assignment of the second event is that of the original event. Both of the new events have zero delay and hence can be mapped to hybrid automata events which by definition have no concept of delay.

In mapping from Bio-PEPA with events to HYPE, the same approach is taken as that used to map Bio-PEPA with events to hybrid automata [26]. Events with a zero delay are treated directly and events with a non-zero delay are split into two events, both with zero delay and hence can be mapped to HYPE instantaneous events. The details of this mapping are now considered.

An event e in Bio-PEPA with a zero delay can be mapped to a HYPE instantaneous event $\underline{\mathrm{e}}$ with event condition $ec(\underline{\mathrm{e}}) = (trigger, species\_assignment)$. However, other assignments which are not related to species quantities involve additional modifications as discussed below.

In the case of events with non-zero delays, it is necessary to define an explicit time variable T, an influence t with $iv(t) = T$, and the subcomponent $Time \stackrel{def}{=} \underline{\mathrm{init}} : (t, 1, 1).Time$ which sets up the appropriate flow for the passage of time. Then for each event e with a non-zero delay, an additional time variable T_e is introduced and two events are generated, namely $ec(\underline{\mathrm{c}}) = (trigger, T'_e = T)$ and $ec(\underline{\mathrm{e}}) = (T = T_e + delay, species\_assignment)$. The timer-setting event $\underline{\mathrm{c}}$ does not need to be included in controllers or subcomponents since they do not need to react to it because it only notes that certain conditions are true and sets a timer to determine when the event that applies the assignments should occur.

The remainder of this section focusses on the events that are involved in event assignments – those that are specified with zero delays and those which are the second event obtained from splitting an event with non-zero delays into two events. These events may have different event assignments which require different treatments. In the case of a modification of a species quantity, only the event needs to be created as described above. However, the case of change in a rate function or change in a constant which is used in a rate function has implications for the subcomponents, controller, and stochastic event relating to that reaction. Volume and change of volume are not considered in this presentation of Bio-PEPA so they are excluded.

Assume that the rate function affected is that of reaction k, and that the events that affect it are $e_1, \ldots, e_n$. For convenience, both a change of rate function and a change of constant as an argument to that rate function will be denoted by a new function f_{k,e_i} where e_i is the event causing the change. For each species S_i involved in reaction k whose rate function is modified by events $e_1, \ldots, e_n$, its subcomponent $SC_{i,k}$ requires additional capabilities which represent flows using these modified functions.

$$
\begin{aligned}
SC_{i,k}(\mathbf{X}) \quad = \quad & \underline{\mathrm{init}} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \underline{\mathrm{stoch}}_k : (\iota_{i,k}, 0, 0).SC_{i,k}(\mathbf{X}) + \\
& \underline{\mathrm{det}}_k : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \sum_{l=1}^{n} \underline{\mathrm{det}}_{k,e_l} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_{k,e_l}(\mathbf{X})).SC_{i,k}(\mathbf{X})
\end{aligned}
$$

Subcomponents do not impose order on the execution of events and this must be done in the subcontroller for reaction k. The following subcontroller switches to different states depending on the event that has happened. Note that only one event $\underline{\text{stoch}}_k$ is required since this simply switches off the flow regardless of what it was.

$$
\begin{aligned}
ConD_k &\stackrel{def}{=} \underline{\text{stoch}}_k.ConS_k + \sum_{l=1}^{n} \underline{\text{e}}_l.ConD_{k,e_l} \\
ConS_k &\stackrel{def}{=} \overline{\text{react}}_k.ConS_k + \underline{\text{det}}_k.ConD_k + \sum_{l=1}^{n} \underline{\text{e}}_l.ConS_{k,e_l} \\
ConD_{k,e_l} &\stackrel{def}{=} \underline{\text{stoch}}_k.ConS_{k,e_l} + \sum_{l=1}^{n} \underline{\text{e}}_l.ConD_{k,e_l} \\
ConS_{k,e_l} &\stackrel{def}{=} \overline{\text{react}}_{k,e_l}.ConS_{k,e_l} + \underline{\text{det}}_{k,e_l}.ConD_{k,e_l} + \sum_{l=1}^{n} \underline{\text{e}}_l.ConS_{k,e_l}
\end{aligned}
$$

Event conditions must be defined for the events that represent the reaction being treated stochastically with the different rate functions.

$$
ec(\overline{\text{react}}_{k,e_l}) = \left(f_{k,e_l}(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k}), \bigwedge_{i \in R_k} X_i' = X_i - \lambda_{i,k} \wedge \bigwedge_{i \in Q_k} X_i' = X_i + \lambda_{i,k} \right)
$$

This mapping of Bio-PEPA with events to stochastic HYPE works even when multiple changes of functions and constants are associated with a single event.

Figure 11 provides an algorithm to add events specified for a Bio-PEPA model to the HYPE representation of that model (without the Bio-PEPA events). To simplify presentation, it will be assumed that an event can either modify species quantities or rate functions and constants for rate functions, but not both at once. So there can be multiple species quantity updates in an event but no rate function modifications and *vice versa*.

The left-hand graph in Figure 12 illustrates a run of the stochastic hybrid model of enzyme kinetics (with reaction thresholds set at 50, and species thresholds all negative) with an addition of more substrate at time 50. This is the results of a Bio-PEPA events of the form $(addS, t = 50, S' = S + 25, 0)$. The addition has an immediate effect (which is modelled deterministically because the reactions are fast). The right-hand graph shows the addition of more enzyme through the events $(addE, t = 50, E' = E + 25, 0)$ and illustrates much less of an effect.

9. Bio-PEPA with delays

Bio-PEPAd [24] is an extension to Bio-PEPA where each reaction can have associated with it a constant duration as well as an exponential duration. The syntax of the Bio-PEPA model remains the same; however, an additional function $\sigma : \mathcal{A} \to \mathbb{R}^+$ which assigns a real value to each reaction.

Constant delays have been introduced in the modelling of biological systems to allow for abstraction in models. This approach is useful when a number of events occur but are modelled as a single event with a constant delay. This may be done either to simplify the model or because the details of the events are not known, although there are measurements of the total duration for all the events. Additionally, a reaction can be viewed as having a stochastic duration based on probabilistic reasoning as well as a constant duration associated with mechanical aspects such as binding.

Deterministic models with delays are defined as delay differential equations [31, 65]. In this approach, values of variables and their derivatives from earlier time points may be used to define the differential equations; whereas in ordinary differential equations, values are used only from the current time. Stochastic approaches are based on modifications of Gillespie's algorithm [16, 5, 4]. The approach taken in Bio-PEPAd is *delay-as-duration* where reactants are removed at the start of the reaction and products added at the end. There are other approaches to modelling delays such as purely-delayed and random-delayed [4, 21, 23, 65].

HYPE does not use delay differential equations but a simple modification to the syntax of influence types could allow this. Currently influence types have the form $I(\mathcal{W})$ and are mapped to functions by $[\![I(\mathcal{W})]\!]$. Consider an influence type $nonlinear(X, Y)$ with $[\![nonlinear(X, Y)]\!] = X \times Y$ which in the context of an ODE is $[\![nonlinear(X, Y)]\!](t) =$

Input: Stochastic HYPE model generated using the algorithm in Figure 5 from a Bio-PEPA model with N species and K reactions of the form $S_1(x_{1,0}) \underset{*}{\bowtie} \ldots \underset{*}{\bowtie} S_N(x_{N,0})$ with thresholds for reactions, $r_1, \ldots, r_k$ and thresholds of reactions in species, $h_{1,1}, \ldots, h_{N,K}$, ε
M Bio-PEPA events of the form (e_l, tr_l, a_l, d_l)

Output: Modified HYPE model

★ Define T as given in the text

★ Let $iv(t) = T$

★ For l in $\{1, \ldots, M\}$, consider (e_l, tr_l, a_l, d_l)

- If $d_l \geq 0$
  - Add a new variable T_l to $\mathbf{X}$
  - Let $ec(\underline{c}_l) = (tr_l, T_l' = T)$
  - Replace event (e_l, tr_l, a_l, d_l) with $(e_l, T = T_l + d_l, a_l, 0)$
- Let $\underline{e}_l$ be an event
- If a_l modifies species quantities
  - $reset(\underline{e}_l)$:= "*true*"
  - For each species modification of the form $S_q = \beta_q$
    - $reset(\underline{e}_l) := reset(\underline{e}_l) +$ " $\wedge\, S_q' = \beta_q$"
  - Let $ec(\underline{e}_l) := (tr_l, reset(\underline{e}_l))$
- Else if a_l modifies rate functions or constants for rate functions
  - For each reaction rate modification giving the form f_{k,e_l}
    - Add " + $\underline{\text{det}}_{e_l} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_{k,e_l}(\mathbf{X})).SC_{i,k}(\mathbf{X})$" to each $SC_{i,k}(\mathbf{X})$
    - Add " + $\underline{\text{det}}_{e_l}.ConS_{e_l}$" to each $ConS_k$
    - Define $ConD_{e_l}$ and $ConS_{e_l}$ as given in the text
    - $act(\overline{\text{react}}_{k,e_l})$:= "f_{k,e_l}"
    - For each reactant in reaction k
      - $act(\overline{\text{react}}_{k,e_l}) := act(\overline{\text{react}}_{k,e_l}) +$ " $\times\, I_{i,k}(\lambda_{i,k})$"
    - Let $ec(\overline{\text{react}}_{k,e_l}) = (act(\overline{\text{react}}_{k,e_l}), reset(\overline{\text{react}}_k))$

Figure 11: Algorithm to add events from a Bio-PEPA model to a HYPE model

$X(t) \times Y(t)$, namely the value of X at the current time t multiplied by the value of Y also at the current time. To specify that the value of X to be used is that of t' time points earlier then, the notation $nonlinear((X, t'), (Y, 0))$ could be used and the resulting function would be $[\![nonlinear(X, Y)]\!](t) = X(t - t') \times Y(t - 0)$. Influence types could then have the general form $I((W_1, t_1), \ldots, (W_q, t_q))$ and the additional time variables would be used in the function definition $[\![I((W_1, t_1), \ldots, (W_q, t_q))]\!]$.

With this extension, Bio-PEPAd could be mapped to HYPE. When a reaction k is to be treated deterministically, then $[\![I((W_1, \sigma(\beta_k)), \ldots, (W_q, \sigma(\beta_k)))]\!]$ would be used and the stochastic treatment would involve the following two events

$$ec(\overline{\text{react}}_k) \quad = \quad \Big(f_k(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k}), \bigwedge_{i \in R_k} X_i' = X_i - \lambda_{i,k} \wedge T_k' = T \Big)$$

$$ec(\underline{\text{finish}}_k) \quad = \quad \Big(T = T_k + \sigma(\beta_k), \bigwedge_{i \in Q_k} X_i' = X_i + \lambda_{i,k} \Big)$$

where T_k' would be a time variable for use by reaction k and T would be the time variable specified in the previous section. The first event would then remove the reactants and take note of the time. The second event would check if the correct time has passed and then add the products. The algorithm presented in Figure 5 could be modified by changing the assignments in the k loop: $reset(\overline{\text{react}}_k)$ would involve storing the current time using T_k but would no longer have the product assignments, a new event $\underline{\text{finish}}_k$ would be introduced with an activation condition that checks if $\sigma(\beta_k)$ time has elapsed and with the product assignments as the resets.

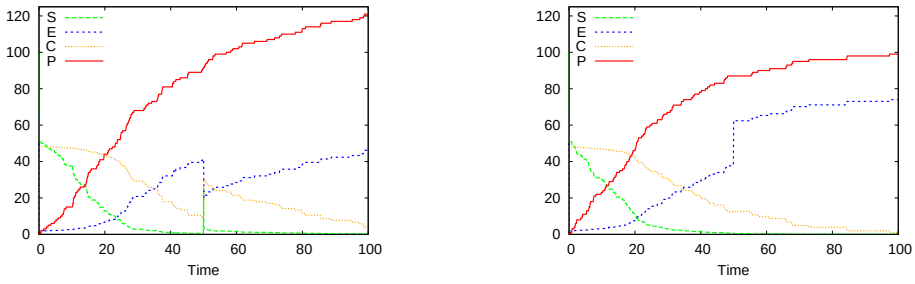

Figure 12: Simulation of the HYPE *EK* model using reaction rate thresholds of 50 and addition of S (left) and simulation of the same model with addition of E (right) with parameters and initial values as in Figure 3 and $\varepsilon = 1$

10. Related work

Previously, PEPA [48] was given hybrid semantics using TDSHA [10]. This is a static approach where actions are determined to be stochastic or deterministic in advance, and this then determines which components are treated continuously and which integrally. To ensure no negative populations, a more complex scheme than proposed here is used to transfer a unit of weight one from continuous components to integral components. Bio-PEPA with events has also been given a hybrid semantics by a mapping to hybrid automata [26]. This mapping only covers deterministic and instantaneous behaviour unlike the approach taken here in mapping to HYPE. An earlier mapping of Bio-PEPA to TDSHA was based on an initial partitioned of species components into discrete and continuous [41] and hence did not consider dynamic switching.

The process algebra stochastic Concurrent Constraint Programming (sCCP) has also been investigated in a hybrid setting using both static [13] and dynamic partitioning [12]. The semantics of a sCCP program is given by a TDSHA, and in the case of dynamic partitioning, in the mapping to the TDSHA each transition for which there is a choice of treatment is split into two transitions, one deterministic and one stochastic with disjoint activation conditions. Note that this is done at the TDSHA level, and not at the process algebra level, in contrast to the approach using HYPE presented here.

Recently, generalised hybrid Petri nets have been introduced for systems biology modelling, particularly for stiff systems [47] together with both static partitioning and dynamic partitioning simulation algorithms. There are many other algorithmic hybrid approaches that use stochastic simulation and a second technique faster than stochastic simulation, both with and without dynamic partitioning of reactions [3, 36, 58]. A particularly interesting approach is the partitioned leaping approach [45, 46]. Depending on the size of the propensities, four different techniques may be used for simulation, ranging from exact stochastic simulation for very small propensities, tau leaping stochastic simulation, solution of Langevin equations to solution of mass action ODEs for larger propensities. None of these approaches provide the language-based approach that is achieved by using process algebra for system description.

11. Conclusion

This paper has presented a mapping from Bio-PEPA to stochastic HYPE which provides a new method of analysis for Bio-PEPA models. The input to the mapping is the Bio-PEPA model together with thresholds to determine dynamically when a reaction should be treated deterministically (as the solution of an ODE system) and when it should be treated stochastically (as an individual reaction occurring in accordance with a functionally-defined exponential rate). The stochastic HYPE model that is obtained provides a clear methodology for how a system of reactions can be modelled and analysed using both deterministic and stochastic simulation combined, giving the advantages of both approaches: speed when deterministic and details when stochastic.

Theoretical results for the mapping show that the resulting model is only slight larger by a constant factor than the Bio-PEPA model. Additionally, setting all thresholds below zero ensures a purely deterministic simulation of the stochastic HYPE model, providing the same results as a deterministic simulation of the original Bio-PEPA model. Letting all thresholds be infinite in the stochastic HYPE model provides a CTMC (after at most $k + 1$ instantaneous

events at time zero) which is the same as the CTMC obtained from the original Bio-PEPA model. A method was presented for finding an estimate for ε, the quantity that separates thresholds. The case study illustrated how behaviour not present in deterministic modelling can be recovered by allowing some reactions to be treated stochastically for some part of the time. This supports related research that proposes similar modelling methodologies to reduce the cost of the modelling but not at the expense of losing detail of interest [46, 13, 22].

Further research involves more case studies as well as understanding the behaviour of models outside of the two extremes and investigated other conditions for switching between treatment beyond those of threshold on reaction rates and species quantities. The advantage of stochastic HYPE as a language is that these can be introduced in the event conditions of the switching events, and require no writing of simulation engines.

Bio-PEPA with events was originally defined to be mapped to hybrid automata (thus only allowing switching between different deterministic behaviour). In this paper, events in Bio-PEPA can be added to the stochastic HYPE model without restricting the type of behaviour it can display and illustrating the modifiability of the stochastic HYPE model. Furthermore, a proposal was presented for extending stochastic HYPE to deal with delay differential equations so that Bio-PEPAd models can also be translated to stochastic HYPE models.

Acknowledgements: The author thanks Jane Hillston for her comments and the reviewers for their useful suggestions on how to improve the paper. This research was funded by EPSRC and partially supported by the EU project QUANTICOL, 600708.

References

[1] Akman, O., Ciocchetta, F., Degasperi, A., Guerriero, M., 2009. Modelling biological clocks with Bio-PEPA: stochasticity and robustness for the Neurospora Crassa circadian network,. In: Proceedings of the 7th International Conference on Computational Models in Systems Biology (CMSB 2009). LNCS 5688. Springer, pp. 52–67.

[2] Akman, O., Guerriero, M., Loewe, L., Troein, C., 2010. Complementary approaches to understanding the plant circadian clock. In: Proceedings of FBTC'10. EPTCS 19.

[3] Alfonsi, A., Cancès, E., Turinici, G., Ventura, B. D., Huisinga, W., 2005. Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. ESAIM Proceedings, 1–13.

[4] Barbuti, R., Caravagna, G., Maggiolo-Schettini, A., Milazzo, P., 2011. Delay stochastic simulation of biological systems: a purely delayed approach. In: Transactions on Computational Systems Biology XIII. LNCS 6575. pp. 61–84.

[5] Barrio, M., Burrage, K., Leier, A., Tian, T., 2006. Oscillatory regulation of hes1: discrete stochastic delay modelling and simulation. PLoS Computational Biology 2 (9), e177.

[6] Blossey, R., Cardelli, L., Phillips, A., 2006. A compositional approach to the stochastic dynamics of gene networks. In: Transactions on Computational Systems Biology IV. LNCS 3939. Springer, pp. 99–122.

[7] Bortolussi, L., Galpin, V., Hillston, J., 2011. HYPE with stochastic events. In: Proceedings of the Ninth Workshop on Quantitative Aspects of Programming Languages (QAPL 2011). EPTCS 57. Saarbrücken, pp. 120–133.

[8] Bortolussi, L., Galpin, V., Hillston, J., 2012. Hybrid performance modelling of opportunistic networks. In: Proceedings of the Tenth Workshop on Quantitative Aspects of Programming Languages (QAPL 2012). EPTCS 85. Tallinn, pp. 106–121.

[9] Bortolussi, L., Galpin, V., Hillston, J., 2013. Stochastic HYPE: Flow-based modelling of stochastic hybrid systems. Submitted for journal publication.

[10] Bortolussi, L., Galpin, V., Hillston, J., Tribastone, M., 2010. Hybrid semantics for PEPA. In: Proceedings of QEST 2010. pp. 181–190.

[11] Bortolussi, L., Policriti, A., Jul. 2008. Hybrid approximation of stochastic process algebras for systems biology. In: IFAC World Congress.

[12] Bortolussi, L., Policriti, A., 2009. Hybrid semantics of stochastic programs with dynamic reconfiguration. In: Proceedings of the Second International Workshop on Computational Models for Cell Processes (COMPMOD 2009). EPTCS 6. pp. 63–76.

[13] Bortolussi, L., Policriti, A., 2010. Hybrid dynamics of stochastic programs. Theoretical Computer Science 411, 2052–2077.

[14] Bortolussi, L., Policriti, A., 2013. (Hybrid) automata and (stochastic) programs: The hybrid automata lattice of a stochastic program. Journal of Logic and Computation 23, 761–798.

[15] Box, G., Draper, N., 1987. Empirical Model-Building and Response Surfaces. Wiley.

[16] Bratsun, D., Volfson, D., Tsimring, L., Hasty, J., 2005. Delay-induced stochastic oscillations in gene regulation. PNAS 102, 1459314598.

[17] Brugnano, L., Mazzia, F., Trigiante, D., 2011. Fifty years of stiffness. In: Recent Advances in Computational and Applied Mathematics. Springer, pp. 1–21.

[18] Cao, Y., Gillespie, D., Petzold, L., 2005. Accelerated stochastic simulation of the stiff enzyme-substrate reaction. Journal of Chemical Physics 123 (14), 144917.

[19] Cao, Y., Gillespie, D., Petzold, L., 2005. Avoiding negative populations in explicit Poisson tau-leaping. Journal of Chemical Physics 123, 054104.

[20] Cao, Y., Gillespie, D., Petzold, L., 2005. The slow-scale stochastic simulation algorithm. Journal of Chemical Physics 122 (1), 14116.

[21] Caravagna, G., 2011. Formal modeling and simulation of biological systems with delays. Ph.D. thesis, University of Pisa.

[22] Caravagna, G., Antoniotti, M., Mauri, G., Graudenzi, A., d'Onofrio, A., 2012. Effects of delayed immune-response in tumor immune-system interplay. In: Proceedings of the First International Workshop on Hybrid Systems and Biology (HSB). EPTCS 92. pp. 106–121.

[23] Caravagna, G., Graudenzi, A., d'Onofrio, A., 2013. Distributed delays in a hybrid model of tumor-immune system interplay. Mathematical Biosciences and Engineering 10, 37–57.
[24] Caravagna, G., Hillston, J., 2012. Bio-PEPAd: A non-markovian extension of Bio-PEPA. Theoretical Computer Science 419, 26 – 49.
[25] Ciocchetta, F., 2008. Bio-PEPA with SBML-like events. In: Proceedings of the Workshop on Computational Models for Cell Processes. pp. 11–22.
[26] Ciocchetta, F., 2009. Bio-PEPA with events. In: Transactions on Computational Systems Biology XI. LNCS 5750. Springer, pp. 45–68.
[27] Ciocchetta, F., Degasperi, A., Heath, J., Hillston, J., 2010. Modelling and analysis of the NF-κB pathway in Bio-PEPA. In: Transactions on Computational Systems Biology XII. LNCS 5945. Springer, pp. 229–262.
[28] Ciocchetta, F., Hillston, J., 2008. Bio-PEPA: a framework for the modelling and analysis of biological systems. Tech. Rep. EDI-INF-RR-1231, School of Informatics, University of Edinburgh.
[29] Ciocchetta, F., Hillston, J., 2008. Calculi for biological systems. In: Formal Methods for Computational Systems Biology, SFM08. LNCS 5016. pp. 265–312.
[30] Ciocchetta, F., Hillston, J., 2009. Bio-PEPA: a framework for the modelling and analysis of biological systems. Theoretical Computer Science 410 (33-34), 3065–3084.
[31] Culshaw, R., Ruan, S., 2000. A delay-differential equation model of HIV infection of CD4+ T-cells. Mathematical Biosciences 165, 27–39.
[32] Curtiss, C., Hirschfelder, J., 1952. Integration of stiff equations. Proceedings of the National Academy of Sciences of the United States of America 38, 235.
[33] Danos, V., Laneve, C., 2004. Formal molecular biology. Theoretical Computer Science 325, 69–110.
[34] Davis, M., 1993. Markov Models and Optimization. Chapman & Hall.
[35] Ellavarason, K., 2008. An automatic mapping from the Systems Biology Markup Language to the Bio-PEPA process algebra. Master's thesis, University of Trento.
[36] Gadgil, C., 2008. Stochastic modeling of biological reactions. Journal of the Indian Institute of Science 88, 45–55.
[37] Galpin, V., 2011. Equivalences for a biological process algebra. Theoretical Computer Science 412, 6058–6082.
[38] Galpin, V., 2012. Modelling trafficking of proteins within the mammalian cell using Bio-PEPA. In: Proceedings of the 10th International Conference on Computational Methods in Systems Biology (CMSB 2012). LNCS 7605. pp. 374–377.
[39] Galpin, V., Bortolussi, L., Hillston, J., 2009. HYPE: a process algebra for compositional flows and emergent behaviour. In: Proceedings of CONCUR 2009. LNCS 5710. Springer, pp. 305–320.
[40] Galpin, V., Bortolussi, L., Hillston, J., 2013. HYPE: Hybrid modelling by composition of flows. Formal Aspects of Computing 25, 503–541.
[41] Gandin, I., 2011. Integrating continuous and discrete dynamic for Bio-PEPA models of biological systems. Master's thesis, University of Trieste.
[42] Gillespie, D., 1997. Exact stochastic simulation of coupled chemical reactions. Journal of Physical Chemistry 81 (25), 2340–2361.
[43] Gillespie, D., 2001. Approximate accelerated stochastic simulation of chemically reacting systems. Journal of Chemical Physics 115, 1716–1733.
[44] Guerriero, M., 2009. Qualitative and quantitative analysis of a Bio-PEPA model of the gp130/JAK/STAT signalling pathway,. In: Transactions on Computational Systems Biology XI. LNCS 5750. Springer, pp. 90–115.
[45] Harris, L., Clancy, P., 2006. A "partitioned leaping" approach for multiscale modeling of chemical reaction dynamics. Journal of Chemical Physics 125, 144107.
[46] Harris, L., Piccirilli, A., Majusiak, E., Clancy, P., 2009. Quantifying stochastic effects in biochemical reaction networks using partitioned leaping. Physical Review E 79, 051906.
[47] Herajy, M., Heiner, M., 2012. Hybrid representation and simulation of stiff biochemical networks. Nonlinear Analysis: Hybrid Systems 6, 942–959.
[48] Hillston, J., 1996. A compositional approach to performance modelling. Cambridge University Press.
[49] Kalantzis, G., 2009. Hybrid stochastic simulations of intracellular reaction-diffusion systems. Computational Biology and Chemistry 33, 205.
[50] Keizer, J., 1987. Statistical thermodynamics of nonequilibrium processes. Springer.
[51] Kummer, U., Olsen, L., Dixon, C., Green, A., Bornberg-Bauer, E., Baier, G., 2000. Switching from simple to complex oscillations in calcium signaling. Biophysical Journal 79, 1188 – 1195.
[52] Kurtz, T., 1972. The relationship between stochastic and deterministic models for chemical reactions. Journal of Chemical Physics 57, 2976–2978.
[53] Kwiatkowska, M., Norman, G., Parker, D., 2011. PRISM 4.0: Verification of probabilistic real-time systems. In: 23rd International Conference on Computer Aided Verification (CAV'11). LNCS 6806. pp. 585–591.
[54] Kwiatkowski, M., Stark, I., 2008. The continuous π-calculus: A process algebra for biochemical modelling. In: Proceedings of CMSB 2008. LNCS 5307. pp. 103–122.
[55] McKane, A., Nagy, J., Newman, T., Stefanini, M., 2007. Amplified biochemical oscillations in cellular systems. Journal of Statistical Physics 128, 165–191.
[56] Michaelis, L., Menten, M., 1913. Die Kinetik der Invertinwirkung. Biochemistry Zeitung 49, 333–369.
[57] Nisbet, R., Gurney, W., 1976. A simple mechanism for population cycles. Nature 263, 319–320.
[58] Pahle, J., 2009. Biochemical simulations: stochastic, approximate stochastic and hybrid approaches. Briefings in Bioinformatics 10, 53–64.
[59] Pedersen, M., Plotkin, G., 2010. A language for biochemical systems: design and formal specification. In: Transactions on Computational Systems Biology XII. LNCS 5945. pp. 77–145.
[60] Priami, C., Quaglia, P., 2004. Beta binders for biological interactions. In: Proceedings of the International Conference on Computational Methods in Systems Biology, (CMSB 2004). LNCS 3082. Springer, pp. 20–33.
[61] Priami, C., Regev, A., Shapiro, E., Silverman, W., 2001. Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information Processing Letters 80, 25–31.
[62] Regev, A., Panina, E., Silverman, W., Cardelli, L., Shapiro, E., 2004. BioAmbients: an abstraction for biological compartments. Theoretical Computer Science 325, 141–167.

[63] Regev, A., Shapiro, E., 2002. Cellular abstractions: Cells as computation. Nature 419, 343.
[64] Samoilov, M., Arkin, A., 2006. Deviant effects in molecular reaction pathways. Nature Biotechnology 24, 1235–1240.
[65] Smith, H., 2011. An Introduction to Delay Differential Equations with Applications to the Life Sciences. Springer.
[66] Tribastone, M., Duguid, A., Gilmore, S., 2009. The PEPA Eclipse Plugin. SIGMETRICS Performance Evaluation Review 36, 28–33.
[67] Troein, C., Corellou, F., Dixon, L., van Ooijen, G., ONeill, J., Bouget, F.-Y., Millar, A., 2011. Multiple light inputs to a simple clock circuit allow complex biological rhythms. The Plant Journal 66, 375–385.
[68] Vasudeva, K., Bhalla, U., 2004. Adaptive stochastic-deterministic chemical kinetic simulations. Bioinformatics 20, 78–84.
[69] Vellela, M., Qian, H., 2007. A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox. Bulletin of Mathematical Biology 69, 1727–1746.
[70] Wolkenhauer, O., Ullah, M., Kolch, W., Cho, K.-H., 2004. Modeling and simulation of intracellular dynamics: choosing an appropriate framework. IEEE Transactions on Nanobioscience 3, 200–207.

Appendix A. Full HYPE model

Consider a well-defined Bio-PEPA system $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, Comp, P\rangle$ with K reactions $\{\beta_1, \ldots, \beta_K\}$, N species, each defined by $S_i \stackrel{def}{=} \sum_{j=1}^{n_i} (\alpha_{i,j}, \kappa_{i,j})\ \mathsf{op}_{i,j}\ S_i$, with $\alpha_{i,j} \in \{\beta_1, \ldots, \beta_K\}$ and $P \stackrel{def}{=} S_1(x_{1,0}) \bowtie_* \ldots \bowtie_* S_N(x_{N,0})$. The following describes the HYPE model $(BP, \mathbf{X}, IN, IT, \mathcal{E}_d, \mathcal{E}_s, \mathcal{A}, ec, iv, EC, ID)$ where IN, IT, $\mathcal{E}_d$, $\mathcal{E}_s$, $\mathcal{A}$, EC and ID are defined implicitly.

Parameters and additional definitions

$$
\begin{array}{rcl}
h_{i,k} & & \text{species thresholds} \\
r_k & & \text{rate thresholds} \\
\varepsilon & \in & (0,1) \text{ to ensure no overlap in event conditions} \\
J_{i,j} & = & \begin{cases} 1 & \text{if } S_i \text{ is a product in reaction } k \\ -1 & \text{if } S_i \text{ is a reactant in reaction } k \end{cases} \\
R_k & = & \{i \mid S_i \text{ is a reactant in reaction } k\} \\
Q_k & = & \{i \mid S_i \text{ is a product in reaction } k\} \\
\lambda_{i,k} & = & \kappa_{i,j} \text{ where } \beta_k = \alpha_{i,j} \\
I_{i,k}(\lambda_{i,k}) & = & \begin{cases} 1 & \text{if } X_i \geqslant \lambda_{i,k} \\ 0 & \text{otherwise} \end{cases}
\end{array}
$$

Variables (one for each species)

$$\mathbf{X} = (X_1, \ldots, X_N)$$

Subcomponents (one for each species/reaction pair where the species is a reactant or product)

$$
\begin{array}{rcl}
SC_{i,k}(\mathbf{X}) & = & \underline{\text{init}} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \\
& & \underline{\text{det}_k} : (\iota_{i,k},\ \lambda_{i,k} \times J_{i,k},\ f_k(\mathbf{X})).SC_{i,k}(\mathbf{X}) + \underline{\text{stoch}_k} : (\iota_{i,k}, 0, 0).SC_{i,k}(\mathbf{X})
\end{array}
$$

Influence mapping (one for each influence)

$$iv(\iota_{i,k}) = X_i$$

Components (one for each reaction)

$$RC_k(\mathbf{X}) \stackrel{def}{=} \bowtie_*{}_{i \in R_k \cup Q_k}\ SC_{i,k}(\mathbf{X})$$

Event conditions (one for each reaction for the last three events)

$$
\begin{array}{rcl}
ec(\underline{\text{init}}) & = & (true,\ \bigwedge_{i=1}^{N} X_i' = x_{i,0}) \\
ec(\underline{\text{stoch}_k}) & = & (\ f_k(\mathbf{X}) \leqslant r_k \vee \bigvee_{i \in R_k} X_i \leqslant h_{i,k},\ true\ ) \\
ec(\underline{\text{det}_k}) & = & (\ f_k(\mathbf{X}) > r_k + \varepsilon \wedge \bigwedge_{i \in R_k} X_i > h_{i,k} + \varepsilon,\ true\ ) \\
ec(\overline{\text{react}_k}) & = & (\ f_k(\mathbf{X}) \cdot \prod_{i \in R_k} I_{i,k}(\lambda_{i,k}),\ \bigwedge_{i \in R_k} X_i' = X_i - \lambda_{i,k} \wedge \bigwedge_{i \in Q_k} X_i' = X_i + \lambda_{i,k}\ )
\end{array}
$$

Controllers : (one for each reaction)

$$
\begin{array}{rcl}
ConD_k & \stackrel{def}{=} & \underline{\text{stoch}_k}.ConS_k \\
ConS_k & \stackrel{def}{=} & \overline{\text{react}_k}.ConS_k + \underline{\text{det}_k}.ConD_k
\end{array}
$$

Controlled system

$$BP \stackrel{def}{=} (\ RC_1(\mathbf{X}) \bowtie_* \ldots \bowtie_* RC_K(\mathbf{X})\ ) \bowtie_* \underline{\text{init}}.(\ ConD_1 \parallel \ldots \parallel ConD_K\ )$$